KING'S College LONDON

Defra Phase 2 urban model evaluation

King's College London
Cambridge Environmental Research Consultants
Imperial College London
Ricardo-AEA

David Carslaw, Helen ApSimon, Sean Beevers, Daniel Brookes, David Carruthers, Sally Cooke, Nutthida Kitwiroon, Tim Oxley, John Stedman and Jenny Stocker

version: 10th October 2013

Summary key points

General points

1. Five urban air quality models have been considered in this report: *ADMS-Urban* from Cambridge Environmental Research Consultants (CERC); *BRUTAL* (Background, Roads and Urban Transport:modelling of Air quality and Limit values) from Imperial College London, *ERG-Toolkit* and the *KCL CMAQ-Urban* from King's College London and the *PCM* (Pollution Climate Mapping) from Ricardo-AEA. These models are all capable of predicting concentrations at the urban scale and some at the regional scale. Predictions were made in London for a base year of 2008.

2. The models were evaluated for NO_x, NO_2, O_3, PM_{10} and $PM_{2.5}$ at a range of site types from kerbside to urban background in London. Comparisons were also made between the models including an assessment of surface areas above/below different concentration thresholds.

3. The evaluation of all models against PM_{10} and $PM_{2.5}$ concentrations is compromised by particle measurement issues. These issues include the variation in particle mass by measurement technique (TEOM vs. FDMS vs. Partisol), the lack of reliable Partisol measurements during 2008 and a general lack of $PM_{2.5}$ measurements. For $PM_{2.5}$, the difference in concentration resulting from measurement technique is similar to the modelled concentration difference between the models.

4. Threshold-type statistics can be very difficult for models to predict well when concentrations are close to the threshold and this is a trait common across all pollutants. Only a small uncertainty in the predicted concentration can have a large effect on the threshold statistic. This behaviour is seen for NO_2 (hours >200 µg m^{-3}), O_3 (days where the maximum rolling 8-hour mean is >100 µg m^{-3}) and PM_{10} (days >50 µg m^{-3}). In addition, calculated exceedance areas are very sensitive to small changes in concentration. While these statistics are relevant from a health-evidence perspective they do not lend themselves to accurate prediction by models.

5. While this work has not specifically considered emission inventories, it is clear there are important differences between the NAEI (the National Atmospheric Emissions Inventory) and LAEI (London Atmospheric Emissions Inventory), particularly with respect to PM emissions e.g. assumptions related to vehicle non-exhaust emissions. Furthermore, it is also clear that there are important differences in the level of background PM_{10} and $PM_{2.5}$ used in the models. Indeed, these issues in many respects account for the largest differences between the models rather than inherent differences in modelling approach *per se*.

6. Models that produce hourly predictions (ADMS-Urban and KCL CMAQ-Urban) can be evaluated much more comprehensively compared with 'annual' models. One of the principal advantages is checking whether these models 'get it right for the right reason'. For example, it is possible to evaluate how well these models predict the hourly NO_x-NO_2 relationship at a specific site, which can build confidence

regarding their ability to also predict how concentrations of NO_2 will change at a particular site for a given reduction in NO_x.

Source apportionment

7. The models produced concentration predictions by different source sectors and these were aggregated by major source categories to determine the contributions made by each source category to total annual mean concentrations.

8. The background (non-London) contribution is very important for PM_{10} and $PM_{2.5}$. There is however a relatively wide variation in this component, which has a strong influence on the calculated surface area above different threshold concentrations. The models use different approaches for calculating the background component e.g. relying on other regional scale model output or deriving components from measurements. Improved consistency between the models should therefore be possible if agreement was reached on the magnitude and composition of the background component.

9. For PM_{10} there are important differences between the models. At roadside sites exhaust emissions contribute 9.1% for ADMS-Urban, 10.9% for BRUTAL, 8.6% for the ERG-Toolkit and 14.7% for the PCM. These results imply that the effectiveness of measures aimed at reducing exhaust PM could vary widely depending on the model used.

10. The most important difference between the models with respect to PM_{10} concentration relates to estimates of non-exhaust emissions. Non-exhaust emissions include tyre and brake wear and particle resuspension. Not all of the models for 2008 provide estimates of all these components i.e. some lack an estimate of road wear/resuspension. In this case the ERG-Toolkit suggests about two to three times the contribution of the other models and a considerable fraction of the overall mean PM_{10} concentration (29.1%) at roadside sites.

11. Many of the differences between the models (in particular for PM_{10} and $PM_{2.5}$) are therefore driven not by the model used but the assumptions concerning the background component and the assumptions in the inventories e.g. the contribution from non-exhaust vehicle emissions. There is therefore considerable scope for minimising the differences between the models by using a consistent set of input data.

Exceedance areas

12. The results suggest the models can provide comparatively large differences in the calculated exceedance areas across London i.e. the surface area above certain concentration thresholds. The models are broadly split into those that provide km^2 estimates of concentrations with roads considered separately (PCM and BRUTAL) and those that model at any geographic scale as a continuum (ADMS-Urban and ERG-Toolkit). Outputs from the latter two models have been down-sampled to 1 km^2 means to allow as consistent a comparison as possible across all models with the data available.

13. In terms of area of NO_2 >40 μg m^{-3} the models provide a wide range of values: ADMS-Urban = 156, BRUTAL = 70, ERG-Toolkit = 295 and PCM = 66 km^2. Furthermore, these exceedance areas are very sensitive to small changes in predicted concentration. For instance, a ±10% increase/decrease in predicted concentration for ADMS-Urban gives a range in exceedance area from 62 to 312 km^2 i.e. the range encompasses the range predicted across all models. The sensitivity of the surface area calculated by the models to small changes in concentration is in part due to the nature of threshold statistics where a small change in concentration close to a threshold can have a large effect on the surface area calculated.

14. For PM_{10} and $PM_{2.5}$ the surface areas above specific thresholds are very sensitive to the model used. Specifically, the principal factor affecting the area calculations is again the assumption concerning the background (non-London) contribution made towards PM_{10} and $PM_{2.5}$. For example, the ≈2 μg m^{-3} difference between ADMS-Urban and the PCM model background concentration results in very different exceedance area estimates: almost 100% of the surface area of London is predicted to be above 20 μg m^{-3} for ADMS-Urban whereas ≈75% of the area of London is predicted to be below 20 μg m^{-3} using the PCM.

15. It can be shown that data presented as km grid square means will tend to underestimate exceedance areas compared with more spatially resolved predictions. This characteristic can be shown by down-sampling the models that provide spatially detailed predictions. Taking ADMS-Urban as an example, the 1 km resolution grid shows that 156 km^2 exceed 40 μg m^{-3} annual mean NO_2. If the grid resolution is reduced to 100 m then it can be shown that 199 km^2 exceed and for a 20 m resolution 192 km^2 exceed.

Effect of scenarios

16. On average the models give a very similar reduction in NO_x due to a scenario that reduces NO_x emissions from road transport by 30% (19–22 μg m^{-3}) based on predictions at monitoring sites. However, the models give a wider range of responses to NO_2 concentrations from 6.4 μg m^{-3} (ADMS-Urban) to 11.1 μg m^{-3} (BRUTAL). The reductions in NO_2 range from 12 to 21%. ADMS-Urban and KCL CMAQ-Urban tend to give a smaller change in NO_2 for a specific change in NO_x concentration and this is most notable for locations with higher concentrations of NO_x. Unlike the other models assessed, ADMS-Urban and KCL CMAQ-Urban consider time-dependent chemical reactions on an hourly basis rather than using empirical NO_x-NO_2-O_3 relationships. These results imply that ADMS-Urban and KCL CMAQ-Urban will tend to predict less of a change in NO_2 for a given intervention that reduces NO_x compared with empirically-based models.

17. A 30% reduction in domestic, commercial and public NO_x emissions reduces NO_2 concentrations at the receptors considered by ≈1.5 μg m^{-3} (3%) for ADMS-Urban, ERG-Toolkit and the PCM but 4.5 μg m^{-3} (9%) for BRUTAL. Important factors here are the release characteristics and immediate vertical dispersion close to the source.

18. Reductions in exhaust PM_{10} emissions of 30% result in a relatively large range in PM_{10} concentration reductions. These reductions will be driven more by inventory

assumptions rather than differences in the modelling approaches used because as noted above reductions in NO_x are consistent across the models. Overall, the models using the LAEI (ADMS-Urban, ERG-Toolkit and KCL CMAQ-Urban) tend to show smaller changes in PM_{10} and $PM_{2.5}$ concentration than those using the NAEI (PCM and BRUTAL).

Analysis of hourly data

19. The models that predict hourly concentrations (ADMS-Urban and KCL CMAQ-Urban) can be scrutinised in much more detail than models that only output annual mean concentrations, which can lead to a better understanding of the underlying reasons why the models behave as they do. The increased temporal output of these models is important with respect to model evaluation.

20. ADMS-Urban and KCL CMAQ-Urban generally capture the non-linear relationship between NO_x and NO_2 very well across a wide range of site types and NO_x concentration. In addition, as noted in the analysis of NO_x reduction scenarios, ADMS-Urban and KCL CMAQ-Urban yield similar reductions in NO_2 concentration for a given reduction in NO_x.

21. Comparison with measured values reveals that there are situations where sources are likely absent or are misspecified in emission inventories highlighting that poor model performance will on many occasions be due to inventory deficiencies rather than model (dispersion/chemistry) limitations.

22. A comparison with a range of air quality metrics shows that ADMS-Urban and KCL CMAQ-Urban perform similarly well. The most important difference in performance is for the prediction of PM concentrations. KCL CMAQ-Urban, which models the background PM component explicitly (rather than using an assumed value based on ambient measurements) tends to underestimate PM_{10} and $PM_{2.5}$ concentrations. However, this underestimate is common among regional-scale models.

23. There are situations where the models fail to capture the complex mixing within street canyons. While it is not expected 'operational'-type models would perform well at specific locations affected by complicated building layouts, analysis can help highlight the implications of such limitations. At Marylebone Road for example, ADMS-Urban and KCL CMAQ-Urban use some of the algorithms from the Danish OSPM (Operational Street Pollution Model) canyon model. However, bivariate polar plots and other analysis shows that as currently used, the modelling approach fails to capture the important reversal of wind flow characteristic of street canyon recirculation. As a consequence, the performance of these models with hourly data is not good. However, these models *do* provide good estimates of annual mean concentrations. It should be noted that exceedances of pollutants such as NO_2 will increasingly be restricted to environments close to roads and therefore it will become increasingly important that models perform well in these locations.

24. A brief analysis using the same techniques above applied to the OSPM model for a street in Denmark does show the model performs well and captures the main

recirculation effects. These results show that revisiting how OSPM is currently implemented in these models could lead to much improved model performance.

25. The analysis of hourly air quality data using either observed meteorology from Heathrow (as used in ADMS-Urban) or from the Weather Research Forecasting (WRF) model results in very similar findings as far as wind speed and direction dependences are concerned. These results suggest that the use of numerical weather prediction models for the estimate of surface meteorological variables may provide a useful means of model evaluation more widely across the UK.

Contents

1 Introduction

1.1 Background

This report summarises the main finding of Phase 2 of the Defra Model Evaluation Exercise. The analysis follows on from Phase 1 where the models were compared against a large number of monitoring sites in London.[1] The report focuses on the use of models that consider urban air pollution, and specifically the prediction of NO_x, NO_2, O_3, PM_{10} and $PM_{2.5}$ concentrations. In Phase 1, the evaluations were focused on comparing simple metrics such as annual means. Phase 2 builds on this analysis and extends Phase 1 to consider a wider range of issues following input from the model groups. The comparisons focus on London for a 2008 base year.

Model evaluations can cover an enormous range of issues and can involve very detailed types of analysis. The main aims of this report are:

1. To quantify the source apportionment by major source categories for NO_x, PM_{10} and $PM_{2.5}$ at measurement sites across London and where possible to better understand the background component of PM predictions.

2. To quantify the surface areas of London above certain threshold concentrations for NO_x, NO_2, PM_{10} and $PM_{2.5}$ to help understand the differences between the models.

3. To quantify the effect of several emission change scenarios focusing on reducing or increasing London emissions for transport and other sources. This aim should help understand how the different model predictions of concentration respond to changes in source emissions.

4. Finally, where models are able to predict hourly concentrations, to consider these predictions in more detail to develop a better understanding of how the models differ.

It should be stressed that while many useful comparisons can be made between the models, there are limits to the analyses that can be undertaken before specific model sensitivity runs are required e.g. to test the importance of emission assumptions on the predictions. In particular the analysis contained in this report focuses on concentration predictions and does not, for example, consider differences in the emission inventories used. Nevertheless, the results contained in this report should provide a good indication of the likely range in model performance from currently used urban modelling approaches. Strictly speaking, the evaluations in this report consider different *model systems*, which comprise inventories, dispersion models and other associated components.

The model evaluation of $PM_{2.5}$ and PM_{10} is compromised by measurement issues. There remains considerable doubt over PM_{10} and $PM_{2.5}$ measurements and this makes the evaluation of models difficult i.e. knowing which measurements are considered to be 'correct' and which should be used as the basis for model evaluation. For example, in the recent AQEG report on $PM_{2.5}$, transects in $PM_{2.5}$ concentrations across London were considered for the PCM and ADMS-Urban in which $PM_{2.5}$ measurements were overlaid

[1] The Phase 1 report is available here: http://uk-air.defra.gov.uk/library/reports?report_id=654.

with Partisol and FDMS measurements (Figure 5.5 in AQEG (2012)[2]). The AQEG work showed that there were differences of up to ≈3 µg m^{-3} between the two instrument types at background locations (for concentrations in the range 10–15 µg m^{-3}). These differences are large and are similar in magnitude to the differences between these two models. However, more recent, long term and high time resolution compositional measurements for many important species will be available through the NERC ClearfLo campaign for 2012. These new measurements could form the basis for a more comprehensive model evaluation for particulate matter.

1.2 Analysis approaches

The analysis has been conducted using R software (R Core Team 2013) and an R package called *openair* (D. C. Carslaw and Ropkins 2012). The document itself embeds R code that generates all the tables, plots and statistics 'on the fly' using the approach of Xie (Xie 2013a; Xie 2013b). This is a *dynamic reporting* approach where the document must be compiled in order to be produced. All the code used to generate the plots, tables etc. is given in Appendix C, which itself is generated directly from the report automatically. There are several reasons for adopting this approach:

- All the analysis is entirely reproducible: every plot, table etc. can be reproduced by anybody with access to the input data. This approach ensures that the process can be made as open and as transparent as possible as every aspect of the analysis can be scrutinised by others.
- All the tools used to produce this report are free and open-source allowing anyone to access them.
- The approach makes it much more efficient to take account of revised data e.g. if a group finds a problem with the model results, new data are re-imported and the document re-compiled and all the analysis is automatically regenerated.
- The model evaluation necessarily can only cover certain aspects of interest. By making all the code available the modellers can develop the analysis further and indeed use it for other situations.

Most of the analysis in this report does not allow for comparisons with observations e.g. the source apportionment estimates from each model. However, some statistical analysis has been undertaken for the models producing hourly output. There are numerous comparisons that can be undertaken when comparing models with measurements. The approach adopted in this report is to keep the comparisons as brief as possible while focusing on those that are considered to be most useful. Several contrasting model evaluation statistics have been calculated that capture different aspects of model performance. In addition the output from the hourly models have been compared with a range of relevant air quality metrics.

In some cases, simple overall summary statistics are presented covering a range of metrics. These summary statistics are defined in Appendix B. To help with interpretation of these statistics summary results have been ranked in a simple way to show the best

[2] Available here http://www.defra.gov.uk/environment/quality/air/air-quality/committees/aqeg/publish/.

performing model first and so on. The ranking is based on the overall performance for the *Coefficient of Efficiency*, *COE* (Legates and McCabe 2012; Legates and McCabe Jr 1999). The *COE* is a simple, good overall indicator of model performance and often other statistics follow the order e.g. the correlation coefficient, r or the Root Mean Square Error, *RMSE*. It is also easy to interpret — see Appendix B. In particular, a value of 1 indicates a perfect model and a value of zero shows that the predictions are no better than taking the mean of the observations. However, there is no single model evaluation metric that can encapsulate the full range of model performance. For this reason, the range of statistics presented aim to provide a good overview of performance.

1.3 The models considered

The models considered in this report briefly described below. Further details relating to these models are provided in the cited references.

1.3.1 ADMS-Urban

The ADMS-Urban model was first developed in 1996 by Cambridge Environmental Research Consultants (McHugh et al. 1997a; McHugh et al. 1997b). ADMS-Urban is an advanced three-dimensional quasi-Gaussian model calculating concentrations hour by hour, nested within a straight-line Lagrangian trajectory model. The model provides predictions on a continuous surface i.e. there is no distinction between 'background' and 'roadside'. Predictions were available on a variable receptor grid with enhanced resolution ($\approx$ metres) close to roads.

ADMS-Urban uses an an 8-reaction atmospheric chemistry scheme based on Generic Reaction Set (NO, NO_2, O_3, VOC) and a simplified scheme for sulphate generation. Specific account is taken of primary NO_2 fractions. Meteorology is based on hourly surface measurements at Heathrow Airport. The boundary conditions for different species are based on hourly measurements from rural sites around London.

For the current study the model produced hourly predictions of all required species including NO_x, NO_2, O_3, PM_{10} and $PM_{2.5}$. Similarly, all scenarios were modelled, based on the London Atmospheric Emissions Inventory (LAEI). Note that ADMS-Urban considers hourly profiles in emissions sources. Elements of the ADMS modelling system have also been used in the ERG-Toolkit, KCL CMAQ-Urban and PCM.

1.3.2 BRUTAL

BRUTAL is the Imperial College Background Road and Urban Transport model of Air quality Limit values model. It is most similar to the PCM model described below in that it produces predictions at background locations at a 1 km^2 resolution and roadside contributions are considered as a separate increment. Longer range modelling out to European scale is used in estimating background concentrations due to sources outside London. BRUTAL uses source-receptor relationships derived from the PPM (Gaussian type) model for area, volume and point sources to estimate PM_{10}, NO_x and NO_2 concentrations. For roadside concentration enhancement factors are applied for street canyon effects according to population density. The meteorological input is based on simple annual wind roses.

BRUTAL was developed as a high resolution (1 km) road transport module based on the road network to be integrated with the UK integrated assessment model, UKIAM,

in order to estimate traffic emissions across the road network and calculate roadside concentrations of air pollutants.

The only chemistry in BRUTAL is for the conversion of NO_x to NO_2. This is handled by an extension of the quadratic relationship with deviations from the photostationary state and allows for variable fractions of primary NO_2.

For the current study the model produced annual predictions of the following species: NO_x, NO_2, PM_{10} and $PM_{2.5}$. Most scenarios were modelled, based on the NAEI.

1.3.3 ERG-Toolkit

The ERG-Toolkit is a semi-empirical model whose dispersion algorithms are based on ADMS (Kelly et al. 2011). The toolkit predicts concentrations using a kernel modelling technique to describe the initial dispersion. The kernel model relates to a set of model concentration fields that are produced using an emissions source of unity: either 1 g s^{-1} (point sources), 1 g m^{-3} s^{-1} (volume sources) or 1 g km^{-1} s^{-1} (road sources). Each kernel is created using hourly meteorological data from Heathrow Airport and hourly emissions profiles to provide the hour of day and day of week emissions variations. The model provides predictions on a continuous surface i.e. there is no distinction between 'background' and 'roadside'.

ERG-Toolkit uses an empirical approach for calculating NO_2 concentrations from total NO_x and for the calculation of O_3.

For the current study the model produced annual predictions for the following species NO_x, NO_2, PM_{10} and $PM_{2.5}$. Similarly, all scenarios were modelled, based on the London Atmospheric Emissions Inventory (LAEI).

1.3.4 KCL CMAQ-Urban

The Weather Research and Forecasting (WRF) meteorological model, the Community Multiscale Air Quality (CMAQ) regional scale model and the Atmospheric Dispersion Modelling System (ADMS) Roads model (version 3) have been coupled to create CMAQ-urban. WRF provides the meteorological input to the KCL CMAQ model which provides air quality predictions starting at a European scale and through a series of model nests which reduce in size, finally to focus on Greater London.

The CMAQ model provides hourly estimates of air quality concentrations across a 3×3 km London grid, and these outputs are coupled with the ADMS model, also driven using WRF meteorology, to provide fine scale hourly air quality predictions. The KCL CMAQ-Urban model gives hourly average concentrations of NO_x, NO, NO_2, O_3 and PM components at 20m×20 m resolution in London using the CB05 chemical mechanism. The near road chemistry was simulated using a simple chemical scheme described in D. C. Carslaw and Beevers (2005). The reaction rates and photo-dissociation rates were taken from the photolysis rate pre-processor (JPROC), part of the CMAQ run, and the time of flight from road sources, estimated each hour, as a concentration weighted average at each receptor location, assuming a straight line between source and receptor and using WRF wind speed at 10 m height. A detailed description of the model can be found in (Beevers et al. 2012).

KCL CMAQ-Urban produced hourly outputs of all species and provided information on the effect of scenarios. No information was provided on surface (exceedance) statistics.

1.3.5 Pollution Climate Mapping

The PCM (Pollution Climate Mapping) model is the model developed by Ricardo-AEA and currently used in the UK for compliance assessment (Grice et al. 2010). The dispersion modelling is carried out using the ADMS model. Long-range transport of PM is derived from the Lagrangian TRACK model. Regional background levels are calculated by interpolating maps of concentrations from rural measurements. A semi-empirical approach is used to calculate near-road concentrations. The roadside increments are modelled by multiplying the adjusted link emission for each link by a 'calibration coefficient'. The model separately considers predictions at background (1 km^2) and along roads. Predictions are for annual mean concentrations only.

The 'total oxidant' approach based on the work of Jenkin (2004a) and Jenkin (2004b) is used to calculate NO_2 concentrations from total NO_x. The meteorology is based on the Met Office surface site at Waddington.

For the current study the model produced annual predictions for the following species: NO_x, NO_2, PM_{10} and $PM_{2.5}$. Similarly, all scenarios were modelled, based on the National Atmospheric Emissions Inventory (NAEI).

1.4 Primary NO_2 assumptions

For predictions of NO_2, the assumptions for the fraction of NO_x emission that is primary NO_2 could be important — particularly from road transport sources. The models assume the following primary NO_2 fractions:

ADMS-Urban Assumes 22% for all major and minor roads, 11.5% for all airport sources and 10.0% for all other sources.

BRUTAL assumes 10% primary NO_2 for all non-traffic sources, and for traffic sources primary NO_2 is based on NAEI assumptions according to vehicle category and local fleet composition.[3] For roadside enhancements this contribution is super-imposed on the primary-secondary mix in the background grid square.

ERG-Toolkit and KCL CMAQ-Urban Assume 21% on average for road vehicle sources but also vary the values depending on the vehicle fleet on each road (from 16.5% to 28.6%). Shipping is 11%, rail 13%, airports 11%, the rest 5%.

PCM For road traffic area sources: Central London 23.3%, Inner London 20.8%, Outer London 18.5%. For non-traffic area sources: Central London 14%, Inner London 12.8%, Outer London 9.3%. For the traffic increment the values range from 15.4 to 26.1%.

It is likely that these assumptions will affect the model performance and the comparison between models. For road traffic sources, all models assume similar values *on average*. However, ERG-Toolkit, KCL CMAQ-Urban and PCM vary the primary NO_2 by road (vehicle type). For some busy roads in central and inner London the difference in these assumptions could be important for roadside predictions.

[3] Murrells, T., Pierce, M. & Passant, N., 2009, An emissions inventory for primary NO_2 and projections for road transport, Briefing note to Defra, NAEI Ref: 48954007.

2 Source apportionment

2.1 Introduction to source apportionment

This section considers the source apportionment of concentrations of NO_x, PM_{10} and $PM_{2.5}$. Each group was asked to provide a breakdown of the predicted concentrations according to major categories in the emission inventories. There is not an exact agreement in source categories assumed mostly due to the differences between the NAEI and LAEI. Nevertheless, it has been possible to identify major source categories to compare. It has not been possible to compare all model outputs on a completely consistent basis due to the way each model calculates source contributions. ADMS-Urban, BRUTAL and the ERG-Toolkit each provide the total London contribution to the major transport categories (e.g. cars, LGVs). However, the PCM model separately considers a 'roadside' and background contributions, with only the former providing information on the vehicle breakdown. It is not possible using the PCM output to estimate the total car, LGV etc. breakdown by vehicle but only a total road transport contribution. However, calculations of the total road contribution have been made.

Details of the sites used in the analysis are given in Table 2.1.

Table 2.1: Details of the sites used in the source apportionment analysis.

site code	easting	northing	site name	site type
BL0	530123	182014	Camden - Bloomsbury	urban background
BX1	551860	176376	Bexley - Slade Green	suburban
CD3	530057	181285	Camden - Shaftesbury Avenue	roadside
CR4	532583	165636	Croydon - George Street	roadside
EA1	517541	180738	Ealing - Ealing Town Hall	urban background
EA2	520304	180054	Ealing - Acton Town Hall	roadside
EN1	533900	195800	Enfield - Bushhill Park	suburban
GR4	543978	174655	Greenwich - Eltham	suburban
HG1	533891	190707	Haringey - Haringey Town Hall	roadside
KC1	524046	181750	Kensington and Chelsea - North Ken	urban background
LB4	531070	175593	Lambeth - Brixton Road	kerbside
LW2	536241	176932	Lewisham - New Cross	roadside
MY1	528125	182016	Westminster - Marylebone Road	kerbside
TD0	515600	170600	Richmond - National Physical Laboratory	suburban
TH1	537509	180867	Tower Hamlets - Poplar	urban background
TH2	535927	182221	Tower Hamlets - Mile End Road	roadside

The data provided by each group allows a comparison to be made for 16 receptors. Note that in the case of the PCM predictions were made at 14 receptors (CD3 and CR4 were absent) and this may affect the comparisons to a small extent. Summaries of these results have been calculated both in terms of absolute concentrations and as percentages. As noted above it has not been possible to estimate the breakdown by vehicle class for all models. However, in the summary tables the sum of cars, LGVs, HGVs and Bus can be used to compare with a total 'roads' contribution from the PCM model. Note that in the following tables 'gas' means the combustion of natural gas and includes domestic, commercial and a very minor contribution from industrial sources.

2.2 Source apportionment for NO_x

Table 2.2 and Table 2.3 show the source apportionment results for concentrations of NO_x. The 'gas' contribution refers to natural gas combustion, which is dominated by domestic and commercial gas use. 'other.tran' is non-road transport sources that include aircraft, rail, ships and off-road vehicle emissions. The 'background' contribution is the assumed mean level of NO_x concentration to which the London sources contribute. The background concentration consists of both non-London UK sources and sources from further afield.

TABLE 2.2: Mean source contributions to NO_x concentrations by model and major source category for 16 sites (μg m^{-3}).

agg.type	variable	ADMS-Urban	BRUTAL	ERG-Toolkit	PCM
background	cars	10.1	6.7	14.0	0.0
background	LGVs	4.3	1.9	4.6	0.0
background	HGVs	7.6	4.5	5.7	0.0
background	Bus	4.7	3.7	6.9	0.0
background	roads	0.0	0.0	0.0	22.6
background	gas	12.6	14.7	13.7	16.3
background	other.tran	7.7	5.4	6.1	4.8
background	background	13.0	12.3	13.5	11.0
background	other	1.3	1.5	0.5	1.0
roadside	cars	45.0	20.1	49.6	0.0
roadside	LGVs	23.4	5.9	19.4	0.0
roadside	HGVs	27.1	13.7	17.6	0.0
roadside	Bus	41.9	11.1	53.6	0.0
roadside	roads	0.0	0.0	0.0	148.4
roadside	gas	17.1	21.0	17.1	19.2
roadside	other.tran	4.0	5.9	3.8	4.7
roadside	background	13.0	11.5	13.5	11.0
roadside	other	1.6	1.8	0.7	1.1

Considering Table 2.2 and Table 2.3 there are several important characteristics of the models. First, there is good agreement in the estimate or assumption of the background contribution, which accounts for ≈20% of the mean NO_x concentration across the 16 receptors at background locations. In terms of absolute values the models are in close agreement with ADMS-Urban = 13.0, BRUTAL = 12.3, ERG-Toolkit = 13.5 and PCM = 11.0 μg m^{-3}. At the roadside sites the background concentration typically accounts for 6 to 13% of the total NO_x.

There are some potentially important differences in the contribution made by different vehicle types, which is best seen in Table 2.3. For example, across the 16 receptors at roadside sites the ERG-Toolkit assumes a lower contribution from HGVs and a higher contribution from buses compared with either ADMS-Urban or BRUTAL. For this reason the models would be expected to respond differently to scenarios that specifically targeted these vehicles. The differences will be mostly due to differences in the emission rates assumed, or the handling of emissions data in general e.g. day of week effects, specific account of queueing.

A more detailed breakdown of the source apportionment of NO_x is shown in Figure 2.1 and Figure A.1 by site. Figure 2.1 shows the absolute concentration breakdown by major source category together with the measured value at each site.

TABLE 2.3: Mean source contributions to NO_x concentrations by model and major source category for 16 sites (%).

type	variable	ADMS.Urban	BRUTAL	ERG.Toolkit	PCM
background	cars	16.5	13.1	21.6	0.0
background	LGVs	7.1	3.7	7.1	0.0
background	HGVs	12.4	8.8	8.8	0.0
background	Bus	7.6	7.3	10.6	0.0
background	roads	0.0	0.0	0.0	40.5
background	gas	20.6	29.1	21.1	29.3
background	other.tran	12.5	10.6	9.3	8.6
background	background	21.2	24.3	20.7	19.7
background	other	2.1	3.0	0.8	1.8
roadside	cars	26.0	22.1	28.3	0.0
roadside	LGVs	13.5	6.4	11.1	0.0
roadside	HGVs	15.6	15.1	10.1	0.0
roadside	Bus	24.2	12.2	30.6	0.0
roadside	roads	0.0	0.0	0.0	80.5
roadside	gas	9.9	23.1	9.8	10.4
roadside	other.tran	2.3	6.5	2.2	2.5
roadside	background	7.5	12.7	7.7	6.0
roadside	other	0.9	2.0	0.4	0.6

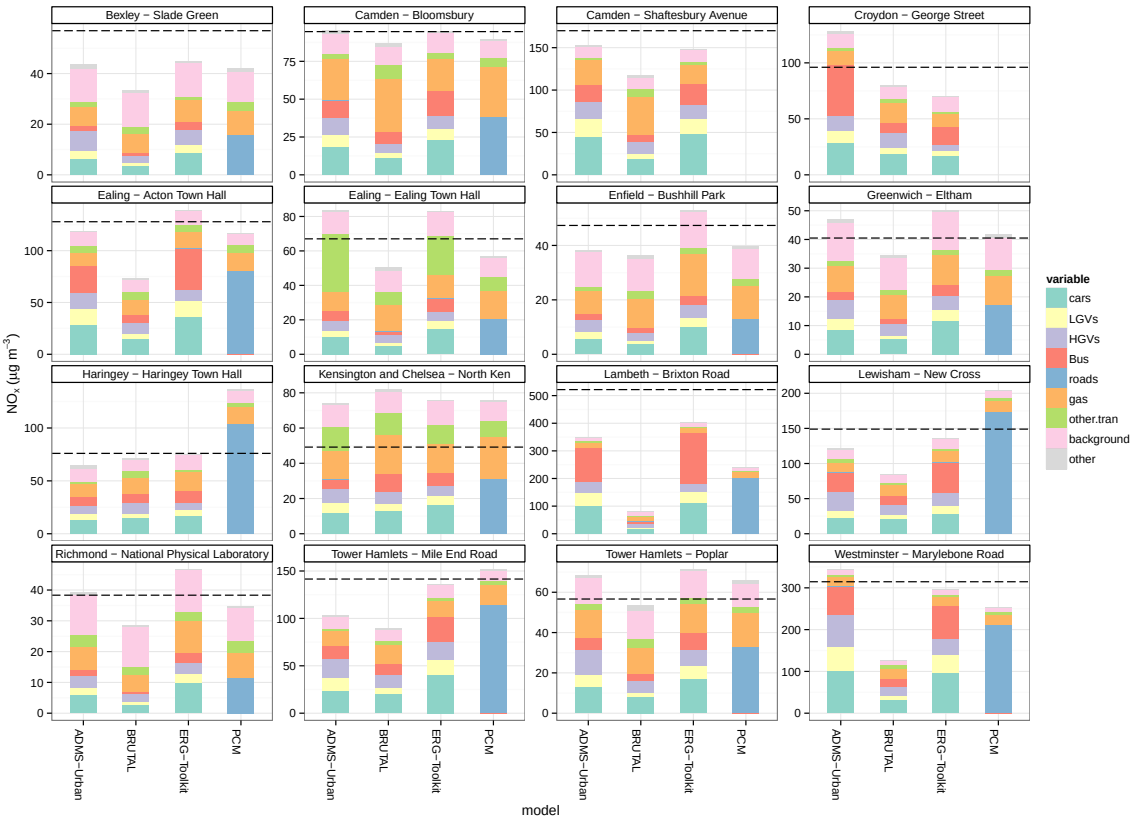

FIGURE 2.1: Source apportionment of NO_x concentrations at 16 sites. The dashed line shows the measured concentration.

2.3 Source apportionment of PM_{10}

A summary of the source apportionment results for PM_{10} by major source category is shown in Table 2.4 and Table 2.5. All models show that the background component contributes considerably to the overall mean concentration across the 16 sites (contributing between ≈69 to 90% of the total at background locations and 55 to 75% at roadside locations).

Figure 2.2 also shows the measured PM_{10} concentration at each site (where available). Note that these concentrations represent values calculated using the KCL *Volatile Correction Model*, VCM. The volatile PM made using nearby FDMS instruments to correct the measurements made by the TEOM. It would also have been useful to compare the predictions with Partisol measurements, but due to problems with filter weight measurements, these were not available for 2008.

In Table 2.4 the vehicle emissions by vehicle type relate to exhaust emissions. The BRUTAL model tends to give around double the contribution for cars compared with ADMS-Urban and the ERG-Toolkit for roadside locations. Exhaust emissions contribute 9.1% for ADMS-Urban, 10.9% for BRUTAL, 8.6% for the ERG-Toolkit and 14.7% for the PCM. These results suggest therefore the models would respond very differently to scenarios that reduce (or increase) exhaust PM emissions.

The principal difference in Table 2.4 and Table 2.5 relates to the estimates of non-exhaust emission estimates. In this case the ERG-Toolkit suggests about two to three times the contribution of the other models and a considerable fraction of the overall mean PM_{10} concentration (29.1%) at roadside sites. Note that ADMS-Urban and ERG-Toolkit non-exhaust estimates include tyre and brake wear and resuspension, the PCM includes it in 'background' and BRUTAL does not include resuspension estimates. The differences observed between the models will be mostly affected by the emission inventories used.

TABLE 2.4: Mean source contributions to PM_{10} concentrations by model and major source category for 16 sites (μg m^{-3}).

agg.type	variable	ADMS-Urban	BRUTAL	ERG-Toolkit	PCM
background	cars	0.2	0.7	0.3	0.0
background	LGVs	0.1	0.3	0.1	0.0
background	HGVs	0.1	0.1	0.2	0.0
background	Bus	0.0	0.1	0.1	0.0
background	roads	0.0	0.0	0.0	1.2
background	gas	0.1	0.1	0.1	0.3
background	other.tran	0.3	0.3	0.2	0.4
background	background	19.7	15.4	16.3	17.3
background	Non.exhaust	0.8	2.3	2.5	0.7
background	other	0.4	3.3	1.7	0.9
roadside	cars	1.0	1.6	1.0	0.0
roadside	LGVs	0.7	0.7	0.3	0.0
roadside	HGVs	0.5	0.3	0.6	0.0
roadside	Bus	0.2	0.2	0.6	0.0
roadside	roads	0.0	0.0	0.0	4.1
roadside	gas	0.2	0.1	0.2	0.3
roadside	other.tran	0.1	0.4	0.1	0.4
roadside	background	19.7	16.1	16.3	18.2
roadside	Non.exhaust	3.4	2.3	8.7	3.9
roadside	other	0.6	3.7	2.0	1.1

TABLE 2.5: Mean source contributions to PM_{10} concentrations by model and major source category for 16 sites (%).

type	variable	ADMS.Urban	BRUTAL	ERG.Toolkit	PCM
background	cars	1.1	2.9	1.6	0.0
background	LGVs	0.7	1.4	0.5	0.0
background	HGVs	0.5	0.7	0.9	0.0
background	Bus	0.1	0.3	0.4	0.0
background	roads	0.0	0.0	0.0	5.7
background	gas	0.6	0.4	0.6	1.3
background	other.tran	1.3	1.3	1.1	2.0
background	background	90.0	68.5	75.5	83.3
background	Non.exhaust	3.9	10.1	11.7	3.3
background	other	1.8	14.5	7.7	4.6
roadside	cars	3.9	6.1	3.4	0.0
roadside	LGVs	2.6	2.8	1.2	0.0
roadside	HGVs	1.8	1.3	1.9	0.0
roadside	Bus	0.8	0.7	2.1	0.0
roadside	roads	0.0	0.0	0.0	14.7
roadside	gas	0.7	0.5	0.6	0.9
roadside	other.tran	0.5	1.6	0.5	1.5
roadside	background	74.8	63.6	54.5	65.3
roadside	Non.exhaust	12.8	8.9	29.1	13.8
roadside	other	2.2	14.4	6.8	3.8

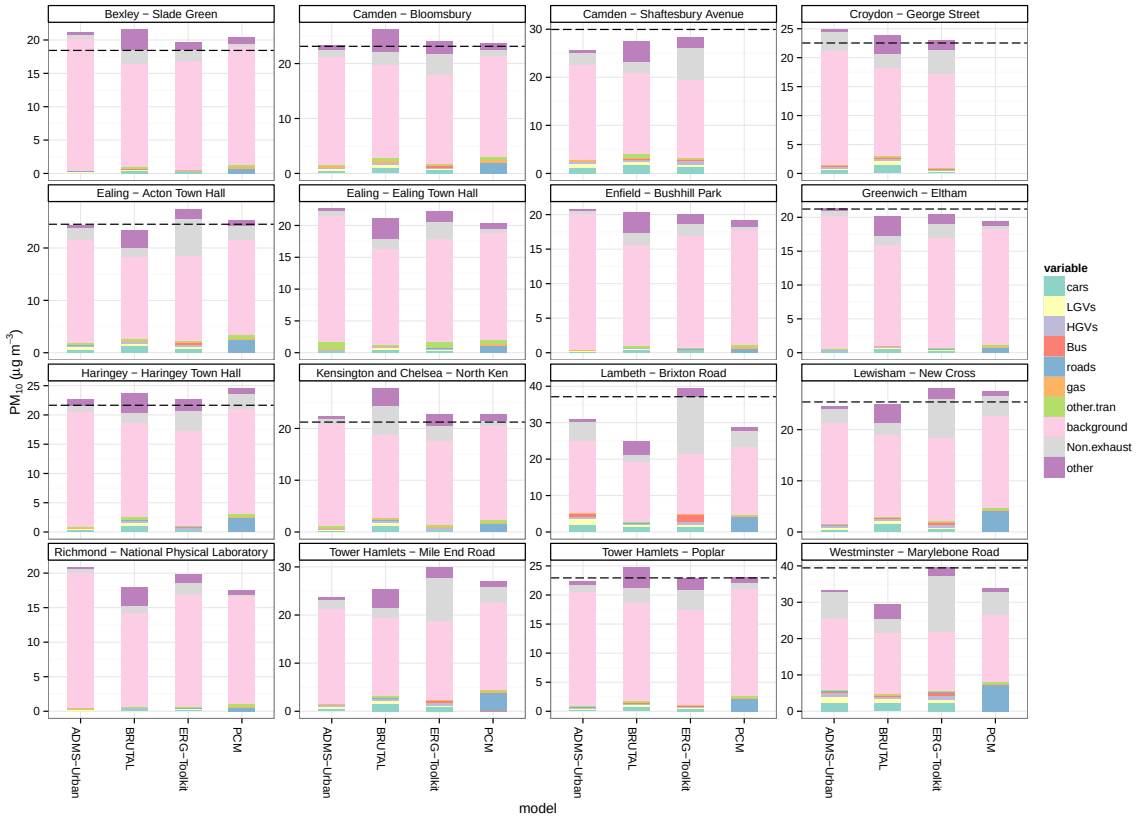

FIGURE 2.2: Source apportionment of PM_{10} concentrations at 16 sites. The dashed line shows the measured concentration.

Given the importance of the background contribution to concentrations of PM_{10}, it is

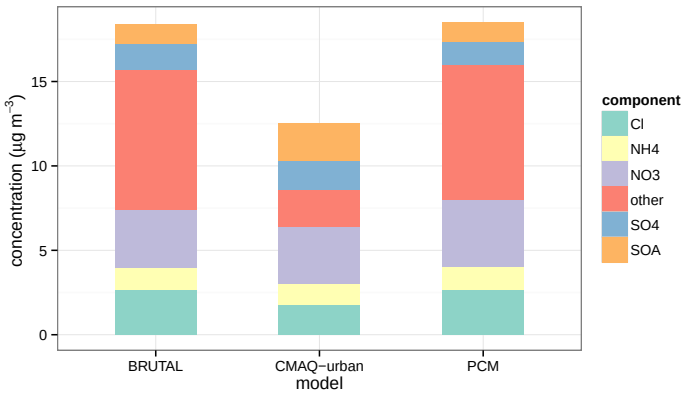

FIGURE 2.3: Source apportionment of mean background PM_{10} concentrations calculated using different models.

worth considering how the models treat this component in more detail. Several of the models provided more information on the disaggregation of the background component including the PCM, BRUTAL and the KCL CMAQ-Urban. For the KCL CMAQ-Urban model the main components of PM_{10} are derived from the CMAQ model directly and vary spatially across London (although only by a small amount) and temporally (hourly values of each component).

Figure 2.3 shows the annual mean breakdown by major component for three of the models: BRUTAL, KCL CMAQ-Urban and PCM. There is generally a good correspondence in the mean values of the major components such as sulphate, ammonium and SOA. The 'other' category contains components such as iron, crustal materials and in the case of the PCM an 'residual'. The principal difference between the models is therefore in these 'other' components. BRUTAL uses the same data as PCM for some of the regional components and this explains the close agreement between them.

2.4 Source apportionment of $PM_{2.5}$

A summary of the source apportionment results for $PM_{2.5}$ are shown in Table 2.6 and Table 2.7, and also in Figure 2.4 and Figure A.3. These results share many of the characteristics of PM_{10}; in particular the large contribution from the background component. The models predict that between 70% to 86% of the $PM_{2.5}$ concentration averaged across the background receptors is due to the background contribution and 58 to 69% at roadside locations. The differences in the background contribution are however important and account for up to 2.3 µg m^{-3} between the models. This difference is important because it is much greater than the vehicle exhaust contribution — on average. Clearly, some level of agreement on what the background concentration 'should' be would help reduce these differences.

Exhaust emissions contribute a relatively small amount of the total $PM_{2.5}$ on average for most models. For example, the average contribution is as follows for roadside locations: ADMS-Urban = 16.5%, BRUTAL = 9.4%, ERG-Toolkit = 14.4% and PCM = 17.3%. The principal difference between the models is for the non-exhaust component where the ERG-Toolkit suggests a mean contribution of 14.7% compared with about 9% for the other models.

TABLE 2.6: Mean source contributions to $PM_{2.5}$ concentrations by model and major source category for 16 sites ($\mu g\ m^{-3}$).

agg.type	variable	ADMS-Urban	BRUTAL	ERG-Toolkit	PCM
background	cars	0.2	0.6	0.4	0.0
background	LGVs	0.1	0.3	0.1	0.0
background	HGVs	0.1	0.1	0.1	0.0
background	Bus	0.0	0.1	0.0	0.0
background	roads	0.0	0.0	0.0	1.4
background	gas	0.1	0.1	0.1	0.3
background	other.tran	0.2	0.2	0.2	0.4
background	background	9.5	9.8	9.2	11.3
background	Non.exhaust	0.3	0.0	0.7	0.5
background	other	0.3	2.9	1.5	0.6
roadside	cars	1.0	0.8	1.1	0.0
roadside	LGVs	0.6	0.4	0.5	0.0
roadside	HGVs	0.4	0.2	0.4	0.0
roadside	Bus	0.2	0.1	0.3	0.0
roadside	roads	0.0	0.0	0.0	3.1
roadside	gas	0.2	0.1	0.2	0.3
roadside	other.tran	0.1	0.3	0.1	0.4
roadside	background	9.5	10.2	9.2	11.8
roadside	Non.exhaust	1.2	0.0	2.3	1.6
roadside	other	0.5	3.1	1.8	0.7

TABLE 2.7: Mean source contributions to $PM_{2.5}$ concentrations by model and major source category for 16 sites (%).

type	variable	ADMS.Urban	BRUTAL	ERG.Toolkit	PCM
background	cars	2.1	4.6	2.9	0.0
background	LGVs	1.3	2.0	1.1	0.0
background	HGVs	1.0	1.0	0.9	0.0
background	Bus	0.2	0.6	0.3	0.0
background	roads	0.0	0.0	0.0	9.6
background	gas	1.1	0.4	1.0	2.2
background	other.tran	2.3	1.6	1.6	3.0
background	background	86.1	69.6	74.6	77.5
background	Non.exhaust	2.9	0.0	5.8	3.4
background	other	3.1	20.2	11.9	4.3
roadside	cars	7.1	5.2	7.0	0.0
roadside	LGVs	4.7	2.4	2.9	0.0
roadside	HGVs	3.2	1.1	2.3	0.0
roadside	Bus	1.5	0.7	2.2	0.0
roadside	roads	0.0	0.0	0.0	17.3
roadside	gas	1.3	0.4	1.0	1.9
roadside	other.tran	0.9	2.1	0.8	2.4
roadside	background	69.1	67.4	57.9	65.6
roadside	Non.exhaust	8.8	0.0	14.7	9.1
roadside	other	3.5	20.6	11.3	3.7

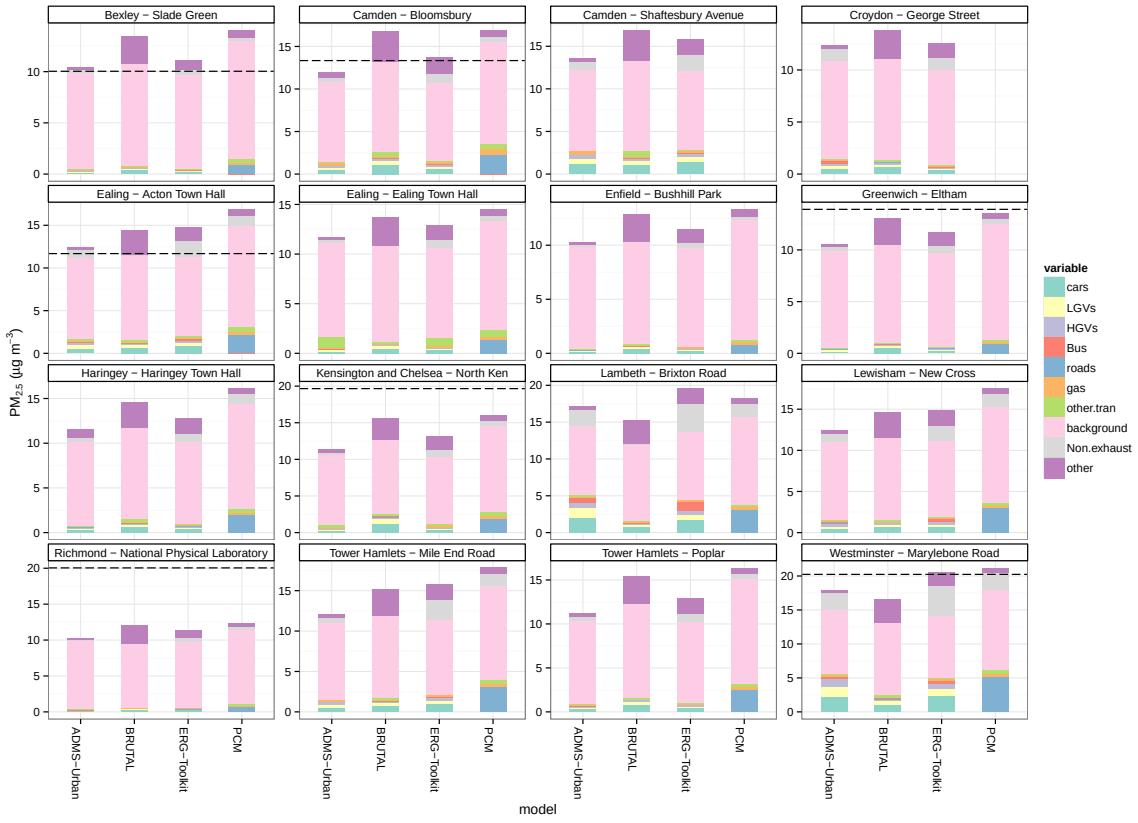

FIGURE 2.4: Source apportionment of $PM_{2.5}$ concentrations at 16 sites. The dashed line shows the measured concentration. Note there are very few sites available for comparison with measurements.

3 Surface predictions

3.1 Exceedance area comparisons

In this section the surface predictions of different species and models are compared. The models provide different output resolutions, so a direct, consistent comparison between them is not possible. The PCM and BRUTAL models predict absolute concentrations on a 1 km^2 grid. In the case of ADMS-Urban the predictions were originally made on a variable grid with more receptors close to roads for example than at background locations. In the case of the ERG-Toolkit predictions were made on a 20 m regular grid. For ADMS-Urban A bi-linear interpolation was carried out to derive concentrations on a 20×20 m regular grid — from which other grid resolutions could be calculated (Akima 2012).

The PCM and BRUTAL models use a 'background' model to estimate concentrations on km^2 basis and therefore do not account for finer spatial details. The PCM however separately estimates length of road network exceeding different concentration thresholds based on a typical single distance from the road, corresponding to the typical distance from the road of a roadside monitor. These data have not been used in the analysis as they would require the other models to define and calculate lengths of road exceeding different limits.

The following comparisons consider the total surface area exceeding different concentration thresholds. To enable a consistent comparison, all model gridded data were averaged to a 1 km^2 basis, although the finer resolution of ADMS-Urban and ERG-Toolkit has been used in several analyses as described below. Furthermore, to ensure the analyses are as consistent as possible, the 2466 km^2 of the LAEI were used to define the area over which the comparisons are made.

3.2 NO_2 exceedance areas

The distribution of NO_2 concentrations on a 1 km^2 basis is shown in Figure 3.1. All models highlight the importance of central London and Heathrow Airport, where concentrations of NO_2 exceed 40 $\mu g\ m^{-3}$. The road sources of ADMS-Urban tend to be more prominent than either the PCM or BRUTAL. However, overall the distributions tend to reflect the importance of road sources. There is also more similarity between ADMS-Urban and ERG-Toolkit and PCM and BRUTAL, perhaps reflecting the use of different inventories.

Higher resolution maps for ADMS-Urban and ERG-Toolkit are shown in 3.2, which emphasise the importance of road sources in controlling the concentration of NO_2. Figure 3.2 shows that ADMS-Urban tends to predict lower concentrations of NO_2 compared with ERG-Toolkit except in the vicinity of Heathrow Airport. The PCM also separately provided an output showing the road link sources superimposed on the gridded background concentrations (Figure 3.3).

Transects of constant latitude or longitude are very useful for exploring the differences between the models. As an example, a transect of constant latitude (at OS 180000 northing) which passes through central London has been analysed. The results, shown in Figure 3.4 highlight some important differences between the models. In terms of concentrations of NO_x, ADMS-Urban and ERG-Toolkit give the highest (and very similar) concentrations across London, but with ERG-Toolkit having higher concentrations in west London. BRUTAL and PCM tend to have lower concentrations, although in the

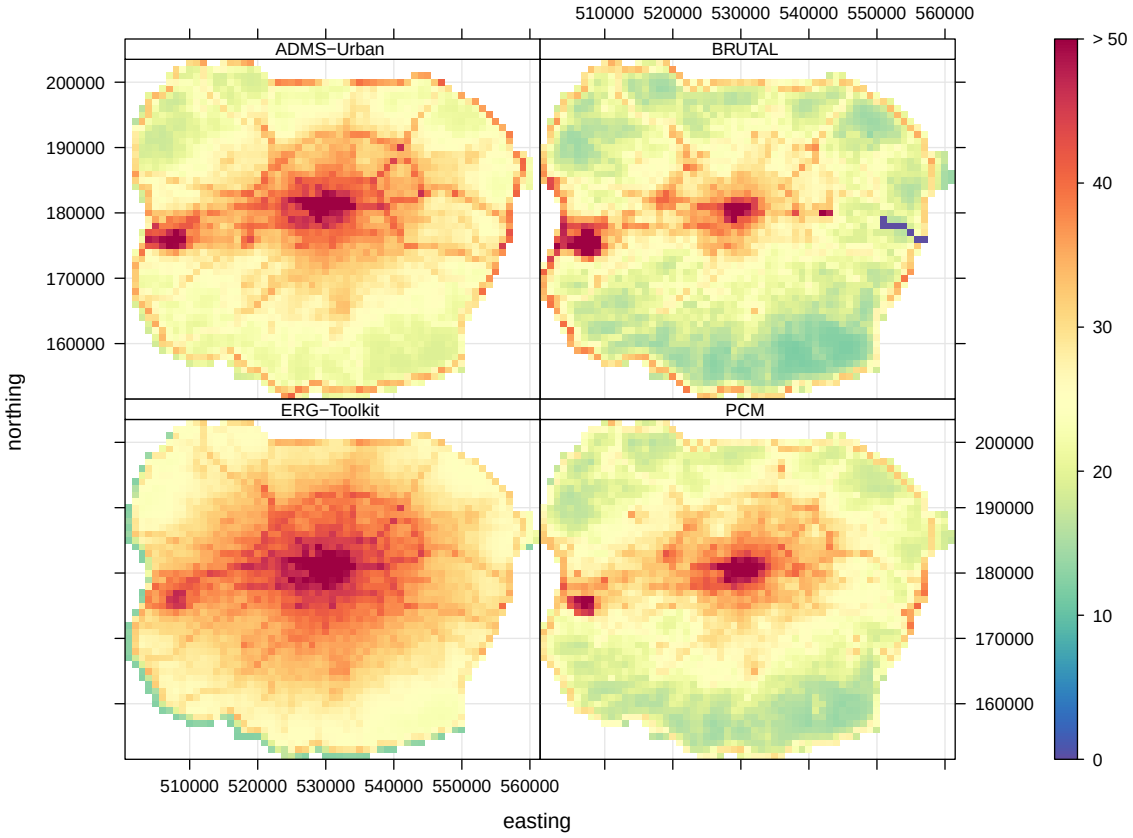

FIGURE 3.1: Surface distributions of annual mean NO_2 concentration ($\mu g\ m^{-3}$).

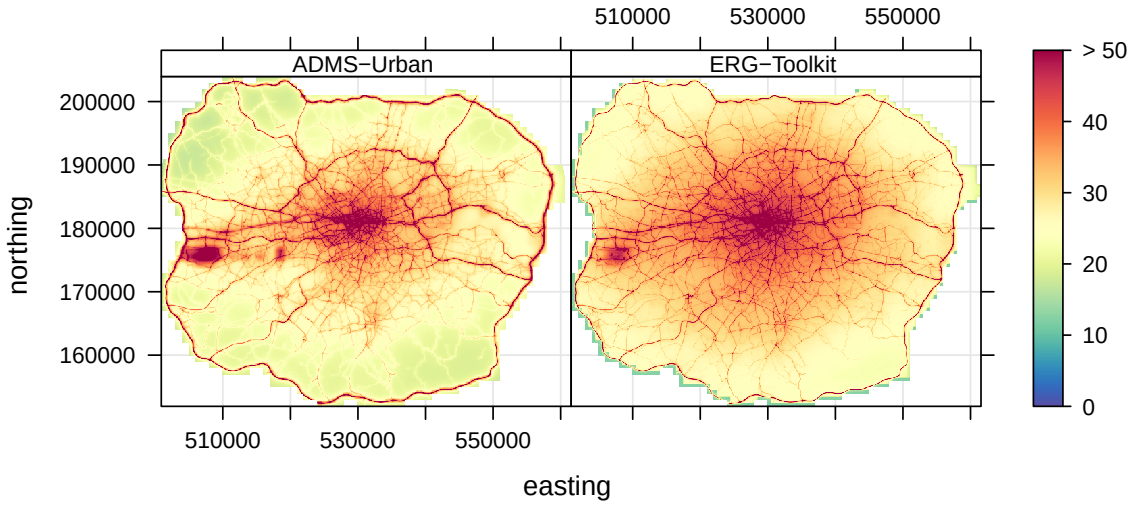

FIGURE 3.2: Surface distributions of annual mean NO_2 concentration ($\mu g\ m^{-3}$) for ADMS-Urban and ERG-Toolkit at a 100 m resolution.

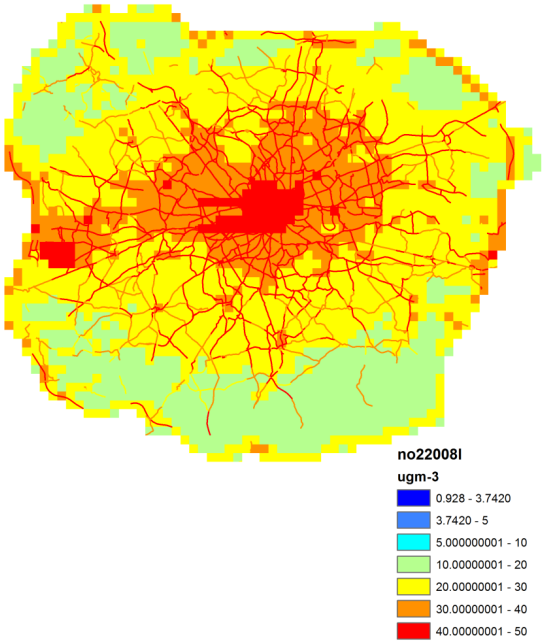

FIGURE 3.3: Spatial distribution of annual mean NO_2 concentrations from the PCM model showing the road link predictions.

east part of central London (around easting = 53000) the PCM is reasonably similar to ADMS-Urban and ERG-Toolkit. There are much larger differences between PCM and BRUTAL compared with ADMS-Urban and ERG-Toolkit for the west part of central/inner London around easting = 525000. These differences are likely related to the differences in the inventories used (LAEI vs. NAEI).

The higher overall NO_x/NO_2 ratio shown by ERG-Toolkit has the effect of producing higher NO_2 predictions across London as shown in the right panel of Figure 3.4. Also, the differences previously noted between ERG-Toolkit and ADMS-Urban in west London become more apparent when considering the transect of NO_2 concentration.

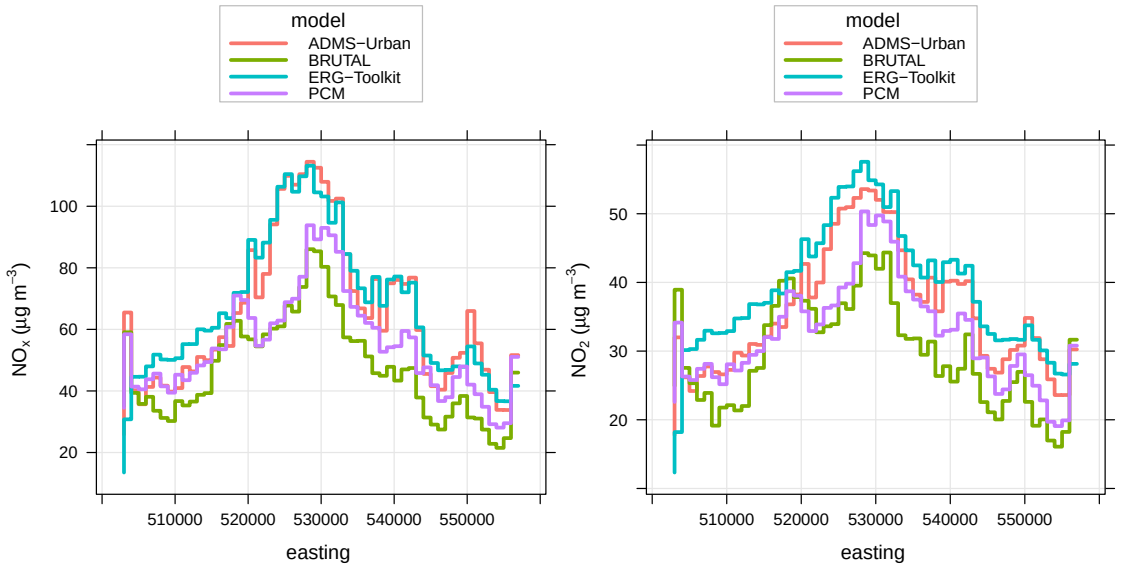

FIGURE 3.4: Transect across 182000 northing for NO_x and NO_2 for concentrations aggregated to 1 km^2.

It is perhaps easier to see the differences between the models by considering Figure 3.5, which shows the cumulative distributions of the NO_x, NO_2 and the NO_2/NO_x ratio. Now it is possible to see the very close similarity of ADMS-Urban and the ERG-Toolkit for NO_x and the lower overall NO_x concentrations predicted by the PCM and BRUTAL. However, the close similarity between ADMS-Urban and ERG-Toolkit for NO_x does not translate to NO_2 where the ERG-Toolkit tends to predict higher concentrations. For NO_2, the ERG-Toolkit distribution has a similar shape to the other models but is offset by $\approx$8 µg m^{-3} compared with ADMS-Urban. This offset is not however related to different estimates of the background concentration (ADMS-Urban and the ERG-Toolkit assume virtually the same background concentration). Given the similarity in NO_x predictions between ADMS-Urban and ERG-Toolkit, the difference in NO_2 predictions is more likely to be due to assumptions related to primary NO_2 and how $NO\text{-}NO_2\text{-}O_3$ chemistry is treated.

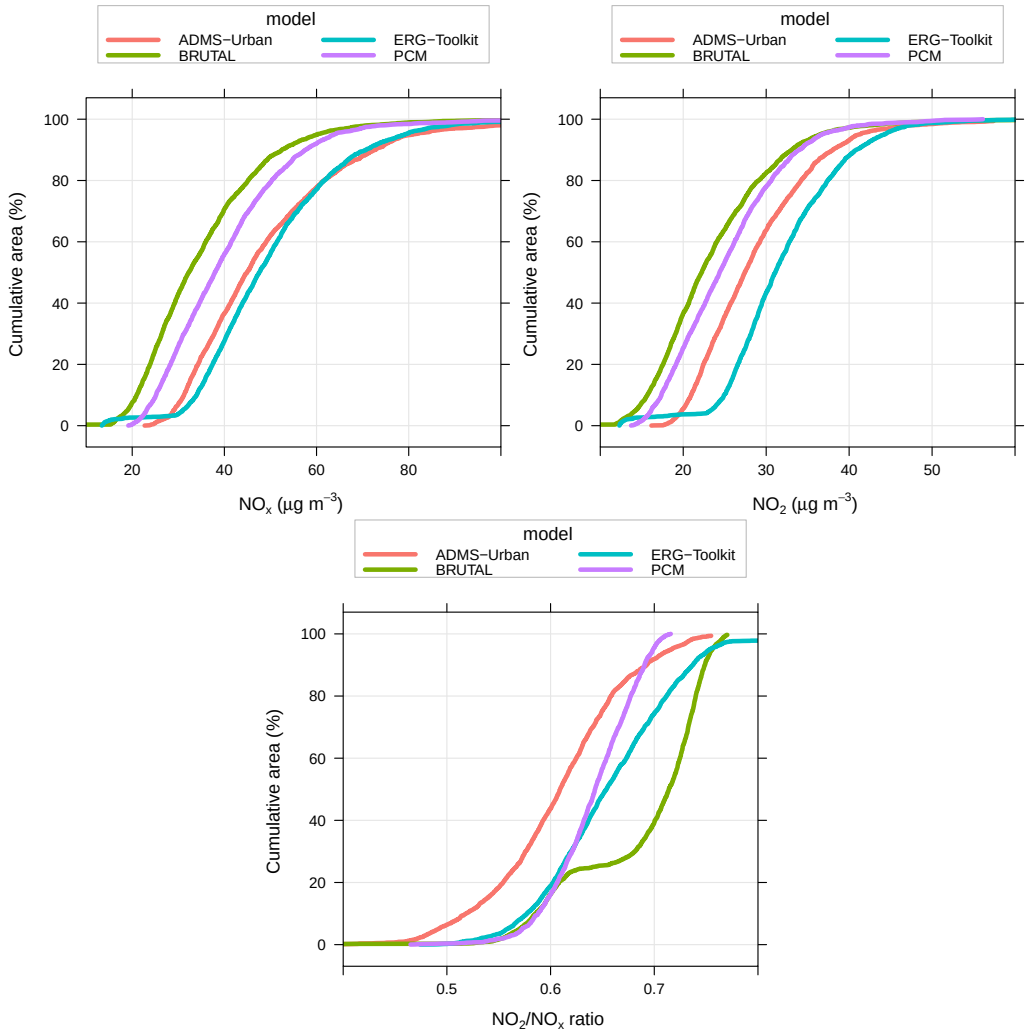

FIGURE 3.5: Cumulative area above different levels of NO_x and NO_2 concentration and the ratio of NO_2/NO_x in London.

These results lead to large differences between the models in their estimate of the surface area of London exceeding different threshold values. For example, it can be shown that the area in km^2 >40 µg m^{-3} for each model is as follows: ADMS-Urban = 156, BRUTAL = 70, ERG-Toolkit = 295 and PCM = 66 km^2.

These results are interesting and important because it is difficult (or impossible) to judge from the comparison of measurements with model predictions whether the use of a particular model would lead to higher or lower numbers of excedenaces across London. For example, in Phase 1 it was shown that the ADMS-Urban and ERG-Toolkit had a very similar performance for annual mean NO_2. However, the ERG-Toolkit model estimates that about twice the area of London would be above the 40 µg m^{-3} Limit Value compared with ADMS-Urban. It should also be noted that these results are a characteristic of threshold statistics, particularly when concentrations are close to a threshold. In these situations even a small change in the predicted value can have a large effect on the exceedance area. Nevertheless, air quality limits are expressed in this way and this behaviour is an intrinsic characteristic of them.

It is also possible to investigate the effect of model resolution on the estimate of area above 40 µg m^{-3} for NO_2 using the models that provide predictions at a high spatial resolution. Taking ADMS-Urban as an example, the 1 km resolution grid shows that

156 km^2 exceed 40 µg m^{-3}. If the grid resolution is reduced to 100 m then it can be shown that 199 km^2 exceed and for a 20 m resolution 192 km^2 exceed. This type of variation is expected because 1 km^2 outputs will tend to reduce the exceedance areas because they are 'averaged-out'. The use of higher resolution outputs therefore tends to increase the exceedance area estimates.

The calculation of surface areas above a threshold such as 40 µg m^{-3} is very sensitive to small variations in the predicted concentration. Given a $\pm$10% variation in predicted NO_2 concentration, the mean and range of the exceedance areas is as follows: ADMS-Urban 156 [62, 312] km^2; BRUTAL 70 [35, 128] km^2; PCM 66 [32, 136] km^2 and ERG-Toolkit 295 [116, 604] km^2. Therefore, given reasonable assumptions about model uncertainty of $\pm$10% it can be shown that the central estimates for each model for the surface area exceeding 40 µg m^{-3} is actually within this range. For example, the range of surface areas exceeding across the models is 66 (PCM) to 295 km^2 (ERG-Toolkit), whereas the $\pm$10% for ADMS-Urban spans a similar range from 62 to 312 km^2.

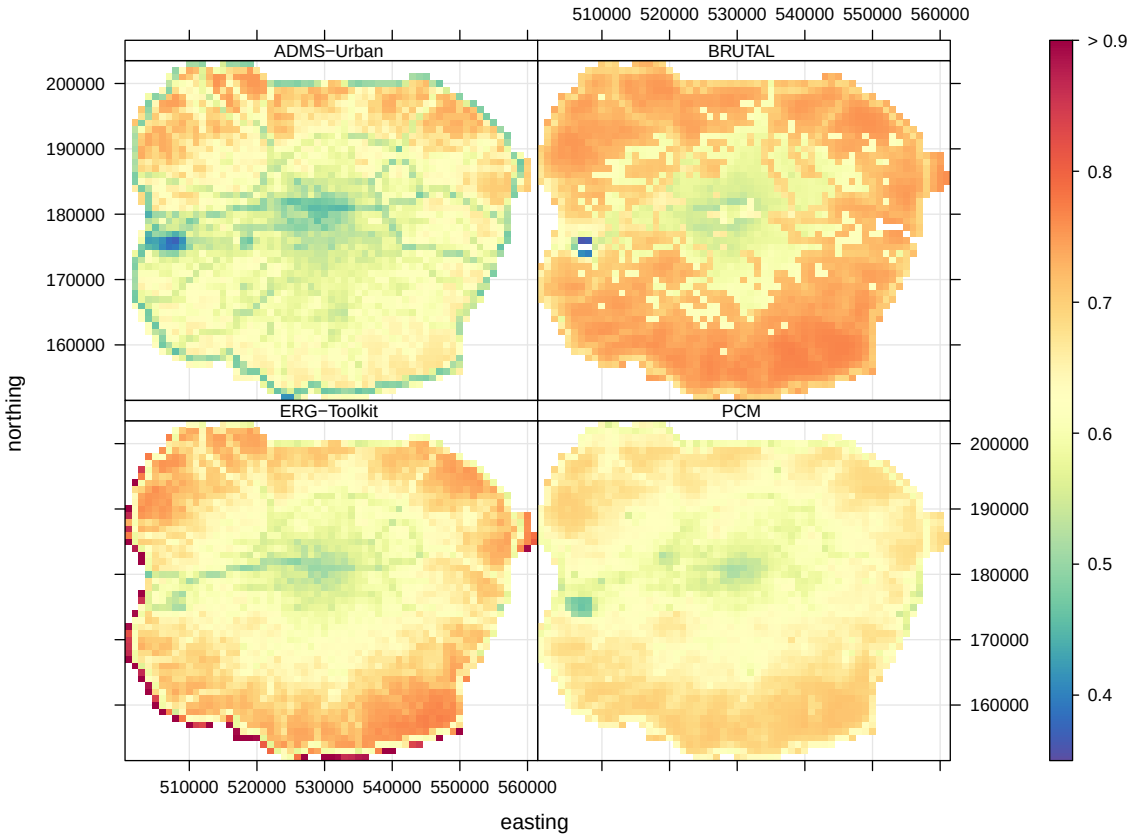

FIGURE 3.6: Surface distributions of annual mean NO_2/NO_x ratio. Note that there are interpolation uncertainties in the ERG-Toolkit results due to the 20 m grids not extending to the edge of the full grid square.

The predictions of NO_x can be used to explore the ratio of NO_2/NO_x spatially. Figure 3.6 shows the NO_2/NO_x ratio across the LAEI grid. It does show some interesting differences between the models. BRUTAL for example tends to have a much higher NO_2/NO_x ratio compared with other models in outer London. ADMS-Urban also shows a lower ratio in central London compared with the other models and a higher ratio in north London compared with south London.

3.3 PM_{10} and $PM_{2.5}$ exceedance areas

The spatial distribution of PM_{10} concentrations for each model is shown in Figure 3.7. An important feature of these maps is the influence of the background concentration of PM_{10} assumed, as this assumption clearly affects the overall concentration field. For example, ADMS-Urban clearly assumes a higher background concentration than other models. The other effect that can be seen is that road sources in BRUTAL are more prominent, suggesting that these emission sources have a greater influence on concentrations of PM_{10} compared with the other models.

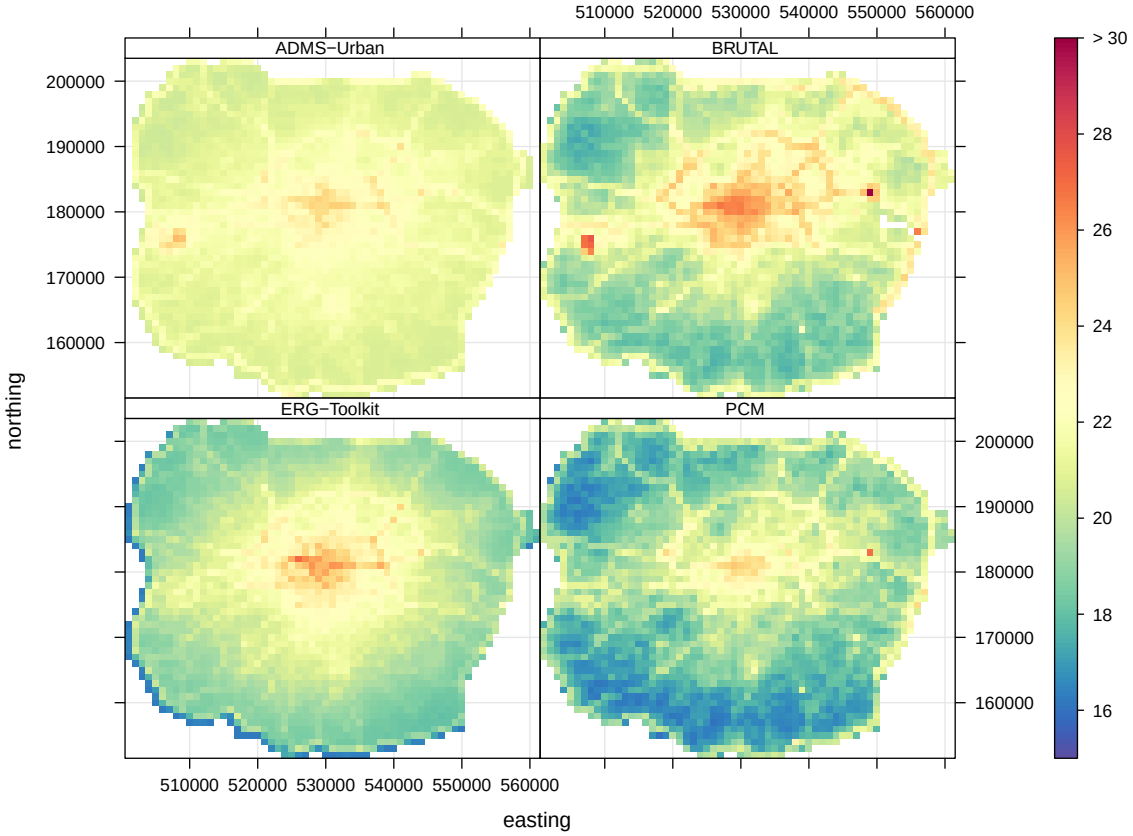

FIGURE 3.7: Surface distributions of annual mean PM_{10} concentration ($\mu g\ m^{-3}$).

East-west transects for PM_{10} and $PM_{2.5}$ are shown in Figure 3.8. It is very clear from this Figure that there are comparatively large differences between the models. ADMS-Urban has a lower urban increment compared with the other models and the PCM tends to estimate lower PM_{10} concentrations in the west part of central/inner London. However, the most significant differences between the models is apparent for concentrations of $PM_{2.5}$, shown in the right panel of Figure 3.8. Here it is clear that there are large differences in the rural background on which the modelled London sources sit. There are also large differences in the location of the highest concentrations. ADMS-Urban not only has a lower concentration profile than the other models, the increment above rural background is also much lower than the other models. Given that NO_x concentrations are reasonably similar across the models, the differences are most likely due to emission source strength differences assumed.

The results for PM_{10} highlight the importance of background concentration assumptions in the models. Figure 3.9 shows that concentrations for all models are compressed into a very narrow range in concentrations. Given the large contribution made by the

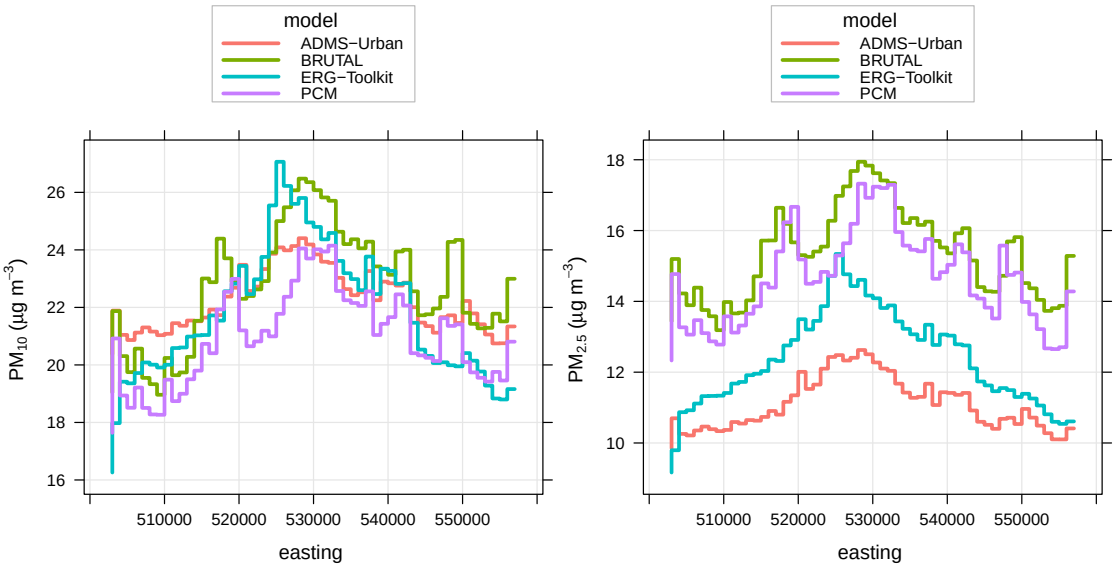

FIGURE 3.8: Transect across 182000 northing for PM_{10} and $PM_{2.5}$ for concentrations aggregated to 1 km^2.

'background' component (as shown in Figure 2.2), even a small variation in that value can have a large effect on the distribution of PM_{10} concentrations. In this case the models tend to follow the assumptions for background contribution (see Table 2.4). For example, in that Table the ADMS-Urban model predicted a higher background concentration than the other models, and this is carried through to Figure 3.9.

The differences between the models are clearly seen in Figure 3.9, which highlights several important features. First, the overall higher background concentration of PM_{10} assumed for ADMS-Urban is apparent. Second, the distribution of PM_{10} for ADMS-Urban is much narrower than the other models which shows that sources in London have less of an affect on PM_{10} concentrations compared with the other models.

The results for $PM_{2.5}$ also highlight the importance of background concentration assumptions in the models. Figure 3.10 shows the spatial distribution of $PM_{2.5}$ across London.

Figure 3.11 shows that concentrations for all models are again compressed into a very narrow range in concentrations. In contrast to the results for PM_{10}, the ADMS-Urban predictions tend to be lower than the other models. This finding is affected mostly by the

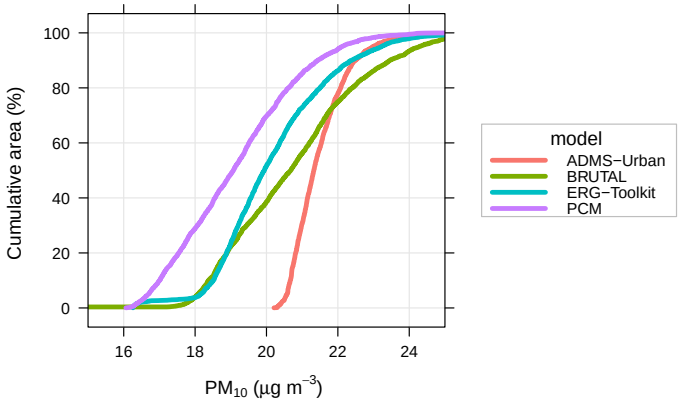

FIGURE 3.9: Cumulative area above different levels of PM_{10} concentration in London.

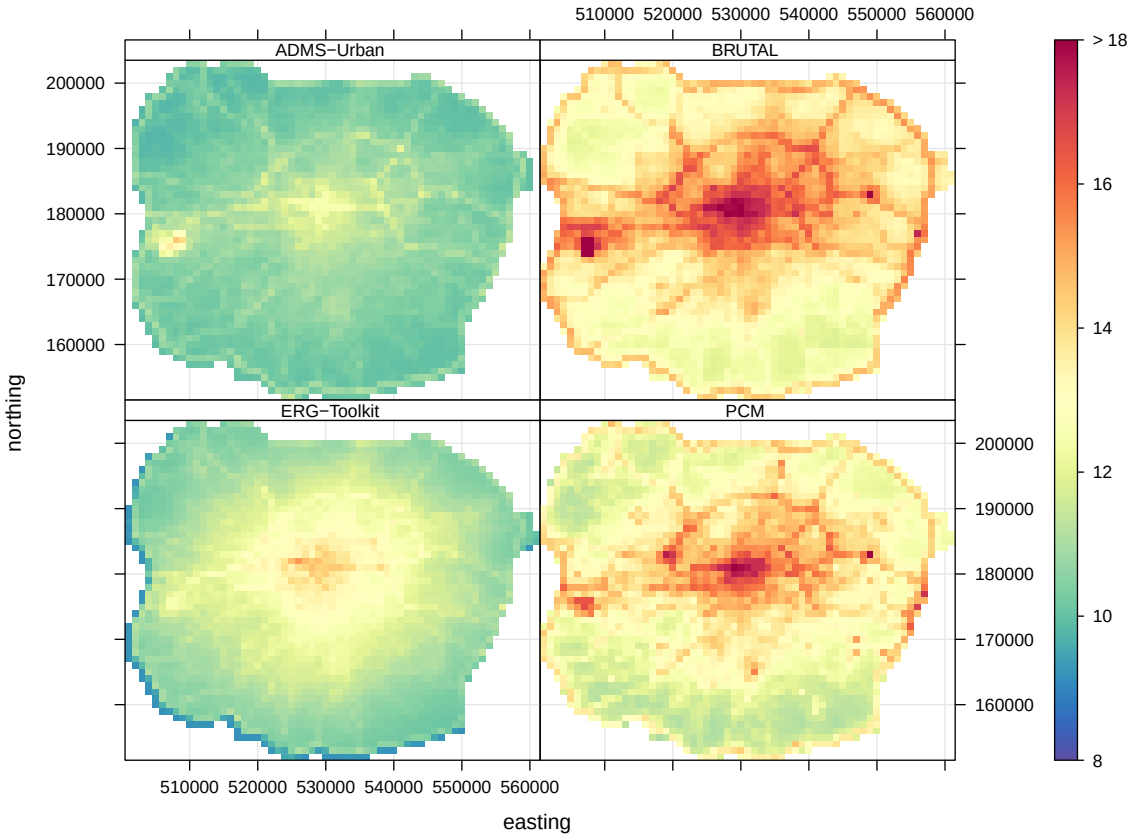

FIGURE 3.10: Surface distributions of annual mean $PM_{2.5}$ concentration (μg m^{-3}).

assumption of background $PM_{2.5}$ concentration assumed. For PM_{10} the ADMS-Urban model predicted higher concentrations than the other models, but this behaviour is reversed for $PM_{2.5}$, as shown in Table 2.6.

The influence of the background assumption on surface area above a concentration threshold is strong and can be most clearly seen in Figure 3.11. For example, for a $PM_{2.5}$ concentration of 12 μg m^{-3} (a typical mean concentration in London, based on the 2008 data), the ADMS-Urban model predicts that almost 100% of the London area will be below this value. The ERG-Toolkit predicts about 85% of the London area will be below this value, while the PCM and BRUTAL models predict very little area of London below this value. The actual differences are very dependent on the background concentration of $PM_{2.5}$ assumed, which as stated before was rather uncertain in 2008 when there were few measurements.

The models could therefore give very different estimates of the area above or below certain concentrations. The extent to which this matters depends very much on the concentration in question: for concentrations around the mean the models are very sensitive to small changes. The key point though is that the surface area exceeding different concentration thresholds is most strongly influenced by the background concentration. Note that the models have different methods of calculating the background component. Therefore, some agreement on the magnitude and composition of this component to ensure consistency across the models would be beneficial.[4]

[4]Note however that the KCL CMAQ-Urban model does explicitly calculate background concentrations but surface area calculations were not available for that model.

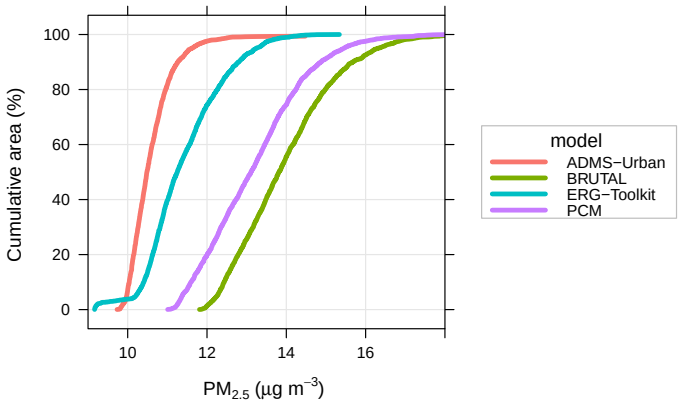

FIGURE 3.11: Cumulative area above different levels of $PM_{2.5}$ concentration in London.

4 Scenarios

4.1 Introduction

This section considers the response of the models to a series of scenarios. The source apportionment analysis in Section 2 provided useful information on the contribution each source category makes to the overall concentration of primary pollutants. In that analysis it was not possible to exactly match the contributions by each model — particularly for road transport emissions where the PCM separately considered roadside and background contributions. While much can be gleaned from the source apportionment analysis, it does not provide any information on how concentrations of secondary component such as NO_2 would respond to source changes. The following scenarios were considered.

S1 30% reduction in road transport *exhaust* emissions of NO_x, PM_{10} and $PM_{2.5}$.

S2 30% increase in road transport *exhaust* emissions of NO_x, PM_{10} and $PM_{2.5}$.

S3 30% reduction in domestic, commercial and public NO_x emissions.

S4 30% reduction in total road transport PM_{10} emissions i.e. exhaust *and* non-exhaust emissions.

These scenarios apply only to emissions in London and no change in background emissions or concentrations has been assumed. In scenarios 1 and 2 only exhaust emissions are considered, which in the case of NO_x and NO_2 could simulate either direct control of exhaust emissions e.g. through after-treatment, or traffic reduction. Scenarios 1 and 2 for PM_{10} and $PM_{2.5}$ would not simulate a reduction in vehicles due to the remaining effect of particle resuspension, which might be expected to be related to total vehicle flows and cannot be controlled through after treatment. It should be stressed however that in the case of PM emissions it is not necessarily the case that a reduction in traffic would result in a commensurate reduction in non-exhaust PM, but the scenario is more to judge the sensitivity of the models to changes in different source components. Note that some groups also predicted the change in PM_{10} and $PM_{2.5}$ for Scenario S3 although these results were not requested from the groups.

The analysis was carried out for all receptors considered in Phase 1.

4.2 Scenario results

A summary of the overall results averaged across all receptors is given in Table 4.1. Note that not all models predicted all scenarios and these are shown as blanks in Table 4.1.

Considering the results for NO_x, the models are in very good agreement across most scenarios. For example, scenario 1 (30% reduction in London road vehicle emissions) covers a narrow range in predicted NO_x reduction from 19.1 to 21.7 µg m^{-3}. Where the models diverge in in scenario 3 (reduction in domestic, commercial emission) where BRUTAL predicts approximately twice the reduction in NO_x compared with the other models. There is an interest in the straightforward question as to what change the models predict due to each scenario, but also whether the changes are consistent. For example, the models may predict similar changes in NO_x concentration but give very different responses for NO_2, which might indicate important differences in how the chemistry is treated in each model.

There is however more of a spread in predicted reduction in NO_2 concentration from 6.4 (ADMS-Urban) to 11.1 µg m^{-3} (BRUTAL) — for scenario S1. The PCM and ERG toolkit reductions are very similar (8.1 and 7.8 µg m^{-3}, respectively), whereas KCL CMAQ-Urban is closer to ADMS-Urban (6.9 µg m^{-3}). These predicted reductions in NO_2 concentration can be compared with a mean measured NO_2 concentration of 52 µg m^{-3}. A 30% reduction in road transport emissions therefore is predicted to reduce NO_2 concentrations by about 12% (ADMS-Urban) to 21% (BRUTAL). The ERG-Toolkit and PCM reductions are 16 and 15% respectively. Note that scenario S2 (30% increase in road transport emissions) leads to similar changes (increases); although they are not as great as the reductions described for scenario S1.

The results for NO_2 can be considered in more detail by retaining the individual site responses. This is important because the response of NO_2 to changes in total NO_x concentration will depend on the concentration of NO_x and the type of site i.e. the extent to which a site is NO_x or O_3 limited. Figure 4.1 shows the response of NO_2 concentrations due to scenario S1. In this plot the response of the sites has been ordered from smallest to greatest change to help with interpretation. There are several important features in this plot that are not apparent from considering Table 4.1. First, the NO_2 response of the models varies considerably by the magnitude of the change in NO_2 concentration. The BRUTAL model gives a much greater reduction in NO_2 across all sites. The PCM, ADMS-Urban, ERG-Toolkit and KCL CMAQ-Urban tend to give similar reductions in NO_2 when the change in concentration is small (i.e. at background sites), but increasingly diverge as the change in concentration increases. There is a discontinuity in the PCM results around a site index of about 54, which could be due to a change from 'background model' to 'roadside model'.

The divergence in the concentration of NO_2 is important because the difference between the models is greatest when the change in concentration is greatest. For example, Figure 4.1 shows that when a change of $\approx$9 µg m^{-3} is predicted by ADMS-Urban and KCL CMAQ-Urban, the other models predict $\approx$15 µg m^{-3}. These are large differences and imply that at sites where the initial NO_x concentration is relatively high, the ADMS-Urban and KCL CMAQ-Urban models predicts much less of a change in the concentration of NO_2 than the other models. The reasons for these differences will be driven by the different methods used to calculate NO_2 concentrations. Models that use empirical relationships to estimate NO_2 from NO_x tend to show more change in NO_2 concentration at higher concentrations of NO_x. Conversely, models that treat the

TABLE 4.1: Effect of the different scenarios by pollutant and model (µg m^{-3}). The base case predictions have been subtracted from each scenario to show the nett effect on concentrations.

pollutant	model	S1	S2	S3	S4
NO2	ADMS-Urban	-6.4	6.1	-1.5	
NO2	BRUTAL	-11.1	10.7	-4.5	
NO2	CMAQ-Urban	-6.9	6.4	-1.5	0.0
NO2	ERG-toolkit	-7.8	7.4	-1.7	
NO2	PCM	-8.1	7.8	-1.6	
NOx	ADMS-Urban	-19.1	19.1	-4.4	
NOx	BRUTAL	-21.7	22.2	-8.7	
NOx	CMAQ-Urban	-18.2	18.3	-4.0	0.1
NOx	ERG-toolkit	-20.4	20.4	-4.3	
NOx	PCM	-19.1	19.1	-4.1	
O3	ADMS-Urban	2.2	-1.9	0.9	
O3	CMAQ-Urban	3.7	-3.1	1.1	-0.0
O3	ERG-toolkit	2.8	-2.4	1.6	
PM10	ADMS-Urban	-0.4	0.4	-0.1	-0.9
PM10	BRUTAL	-1.2	1.7	-0.2	-1.9
PM10	CMAQ-Urban				-0.9
PM10	ERG-toolkit	-0.4	0.4	-0.0	-2.0
PM10	PCM	-0.7	0.7		-1.2
PM25	ADMS-Urban	-0.3	0.3	-0.1	-0.5
PM25	CMAQ-Urban				-0.6
PM25	ERG-toolkit	-0.4	0.4	-0.0	-0.8
PM25	PCM	-0.6	0.6		-0.9

chemistry more explicitly (ADMS-Urban and KCL CMAQ-Urban) show less response.

Hourly predictions for two of the models (ADMS-Urban and KCL CMAQ-Urban) are considered in more detail in Section 5.4.

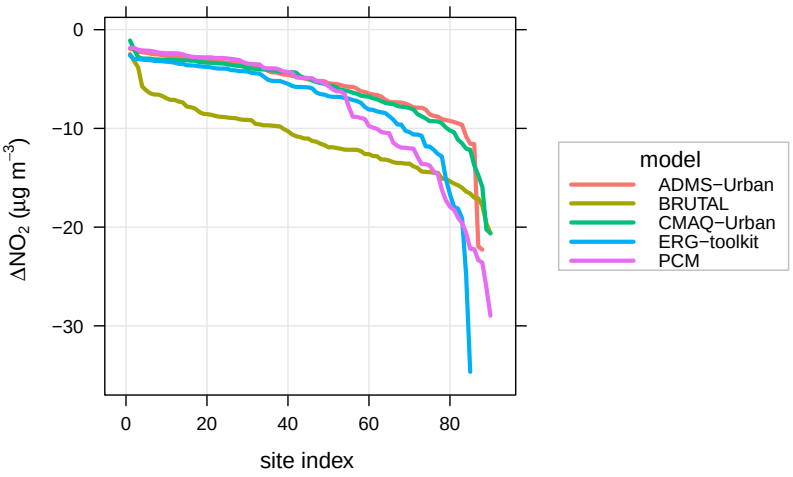

FIGURE 4.1: Effect on concentrations of NO_2 ordered by site for Scenario S1.

Clearly the changes described above for NO_2 and shown in Figure 4.1 ignores the correponding changes in NO_x concentration. Nevertheless, the absolute changes in NO_2 highlight what each model believes the change in NO_2 would be for a particular

scenario. More information can be gained by considering the corresponding change in NO_x concentration, which is shown in Figure 4.2. Now it is possible to see for a particular change in NO_x there is generally less change in NO_2 for the ADMS-Urban model, whereas the PCM and ERG-Toolkit give similar reponses. The BRUTAL model predicts much greater reductions in NO_2 for a particular change in NO_x compared with the other models. These show that the changes in NO_2 concentration seen in Figure 4.1 for the PCM and ERG-Toolkit are in part due to differences in the predicted change in NO_x rather than the relationship between NO_x and NO_2 used.

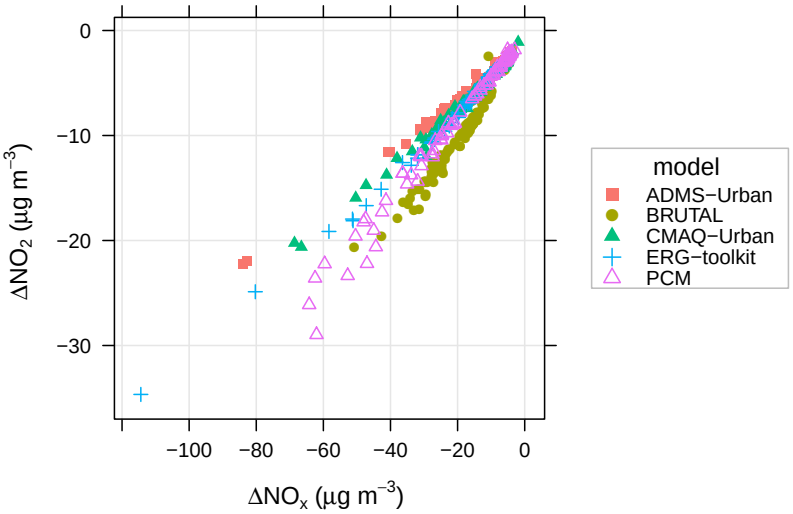

FIGURE 4.2: Change in NO_x and corresponding change in NO_2 concentration ordered by site for Scenario S1.

The changes in PM_{10} and $PM_{2.5}$ are consistent with the transects considered in Section 3.3. For a 30% reduction in exhaust PM, ADMS-Urban and ERG-Toolkit on average predict a similar reduction in PM_{10} of $\approx$0.4 μg m^{-3}. This reduction is much less than suggested by the PCM model of 0.7 μg m^{-3}. Clearly these differences would have important implications for judging the efficacy of emission reduction scenarios. There is a similar situation for $PM_{2.5}$. For scenario 4 the ERG-Toolkit model gives the greatest reduction in PM_{10}, which as seen in Section 2.3 is due the larger contribution assumed from non-exhaust PM. Taken as a whole, the differences between the PM models for *local* emission reduction scenarios will be driven more by the emission inventory data differences rather than the models themselves.

5 Analysis of hourly results

5.1 Introduction

Two of the models were able to provide hourly predictions: ADMS-Urban and the KCL CMAQ-Urban. The prediction of hourly concentrations is very challenging for urban models for several reasons. First, the emissions may not be sufficiently detailed or accurate to adequately capture hourly variations by source. Second, dispersion in urban areas is highly complex and many monitoring sites are affected by complex mixing around buildings that are not represented in models — or are only simply parametrised.

Nevertheless, there are several comparisons with measurements that can be made that can help understand some of the reasons for the differences between models.

The following tables provide summaries of the hourly models for NO_x, NO_2, PM_{10} and $PM_{2.5}$. For NO_x and NO_2 concentrations the performance of the models is very similar; although KCL CMAQ-Urban tends to underestimate NO_x more than ADMS-Urban and ADMS-Urban tends to underestimate NO_2 more. The KCL CMAQ-Urban model uses the WRF model to provide the meteorological data as input to the urban model. The ADMS-Urban model uses the Met Office surface meteorological site at Heathrow Airport. As discussed in the Phase 2 Regional Evaluation report (D. Carslaw et al. 2013), the regional models in general do very well in predicting surface meteorological variables such as temperature, wind speed and wind direction. However, as shown in D. Carslaw et al. (2013) predictions of other important variables such as boundary layer height are much more variable between different models, which will be due to the different planetary boundary layer (PBL) schemes used. Nevertheless ADMS-Urban and KCL CMAQ-Urban produce very similar predictions given the different meteorological data used as input.

Table 5.1: Performance of ADMS-Urban and KCL CMAQ-Urban for NO_x concentrations.

variable	ADMS-Urban	KCL-CMAQ-urban
n	127508.00	133146.00
FAC2	0.73	0.72
MB	-12.29	-24.98
MGE	60.43	61.66
NMB	-0.10	-0.20
NMGE	0.49	0.49
RMSE	115.77	122.23
r	0.73	0.71
COE	0.43	0.43

Table 5.2: Performance of ADMS-Urban and KCL CMAQ-Urban for NO_2 concentrations.

variable	ADMS-Urban	KCL-CMAQ-urban
n	127477.00	133115.00
FAC2	0.81	0.81
MB	-9.84	-5.36
MGE	23.41	23.53
NMB	-0.17	-0.09
NMGE	0.39	0.39
RMSE	43.45	43.68
r	0.71	0.69
COE	0.38	0.38

The results for PM_{10} and $PM_{2.5}$ are very different to those for NO_x and NO_2. The principal difference is that the KCL CMAQ-Urban model underestimates PM_{10} and $PM_{2.5}$ to a much greater extent than ADMS-Urban. There are however considerable differences in how the background component is calculated in each case, which dominates overall concentrations. In the case of ADMS-Urban the background concentration

Table 5.3: Performance of ADMS-Urban and KCL CMAQ-Urban for O_3 concentrations.

variable	ADMS-Urban	KCL-CMAQ-urban
n	73707.00	76937.00
FAC2	0.67	0.65
MB	-0.34	1.99
MGE	11.87	12.75
NMB	-0.01	0.06
NMGE	0.34	0.37
RMSE	16.49	17.20
r	0.79	0.79
COE	0.46	0.42

is based on ambient measurements whereas in the KCL CMAQ-urban model these components are calculated explicitly. Furthermore, it is also known that regional scale models tend to underestimate most PM components.

Table 5.4: Performance of ADMS-Urban and KCL CMAQ-Urban for PM_{10} concentrations.

variable	ADMS-Urban	KCL-CMAQ-urban
n	91735.00	95973.00
FAC2	0.88	0.64
MB	-0.48	-8.68
MGE	8.35	11.09
NMB	-0.02	-0.34
NMGE	0.33	0.43
RMSE	13.12	15.91
r	0.63	0.62
COE	0.30	0.10

Table 5.5: Performance of ADMS-Urban and KCL CMAQ-Urban for $PM_{2.5}$ concentrations.

variable	ADMS-Urban	KCL-CMAQ-urban
n	33588.00	34796.00
FAC2	0.92	0.62
MB	-1.06	-3.77
MGE	4.34	6.41
NMB	-0.07	-0.26
NMGE	0.31	0.45
RMSE	6.70	8.58
r	0.68	0.61
COE	0.33	0.01

5.2 Comparsions with common air quality metrics

The calculation of hourly predictions from ADMS-Urban and KCL CMAQ-Urban allows the performance of the models to be compared against a wide range of air quality

metrics — and thus allows for a more in-depth analysis of the results. In this section some basic comparisons are made with common air quality metrics for each pollutant.

Figure 5.1 summarises the comparison of ADMS-Urban and KCL CMAQ-Urban for three key NO_2 metrics. For the annual mean NO_2 concentration, both models perform similarly well, with almost all NO_2 predictions within a factor of two. Both models struggle to predict the Lambeth site well where observed NO_2 concentrations are very high. Both models also perform well for the higher concentration metric of the maximum daily mean concentration, shown as the middle plot in Figure 5.1. A more limited evaluation of the number of hours where NO_2 exceeds 200 µg m^{-3} (because very few sites have exceedances) shows that this is a difficult metric for the models to predict well.

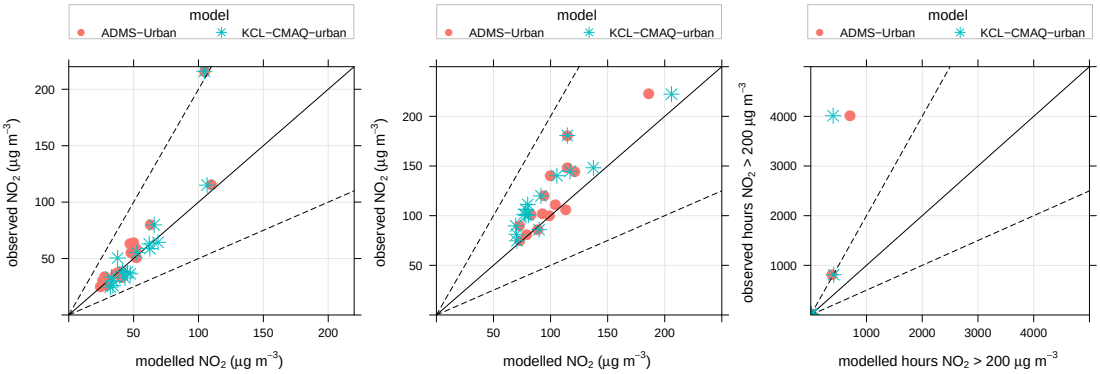

FIGURE 5.1: Comparison of modelled and observed annual mean, maximum daily mean and number of hours >200 µg m^{-3} for NO_2 concentrations for ADMS-Urban and KCL CMAQ-Urban.

The analysis reveals that annual mean O_3 concentrations are predicted well by both models, shown in the left hand plot of Figure 5.2. Both models comfortably predict O_3 concentrations within a factor of two for all sites. ADMS-Urban does rather better at predicting the maximum daily rolling 8-hour mean O_3 concentration (middle plot), where KCL CMAQ-Urban has a tendency to overestimate concentrations. A much tougher statistic for both models is the number of days where rolling 8-hour mean concentrations are greater than 100 µg m^{-3} (right-hand plot). Interestingly in this case, KCL CMAQ-Urban does better than ADMS-Urban, although both models tend to overestimate the total number of days. The performance of the models is therefore dependent on some extent on the metric being considered, but across a wider range of metrics ADMS-Urban tends to have a better performance than KCL CMAQ-Urban.

Two PM_{10} metrics are shown in Figure 5.3: the annual mean PM_{10} concentration and the number of days where PM_{10} is >50 µg m^{-3}. As noted previously the KCL CMAQ-Urban model models the PM_{10} concentrations explicitly and tends to underestimate concentrations, and this is the behaviour seen in Figure 5.3. It is interesting to note however that the KCL CMAQ-Urban and the ADMS-Urban results are similar apart from an offset of about 5 µg m^{-3}. The performance of both models for the London increment is similar, but the regional component is clearly different in each case. In terms of the number of days where PM_{10} is >50 µg m^{-3}, ADMS-Urban again performs better overall, although both models tend to underestimate the total number of days.

Threshold-type statistics can be very difficult for models to predict well when concentrations are close to the threshold and this is a trait common across all pollutants. Only a small uncertainty in the predicted concentration can have a large effect on the threshold statistic. This behaviour is seen for NO_2 (hours >200 µg m^{-3}), O_3 (days where the maximum rolling 8-hour mean is >100 µg m^{-3}) and PM_{10} (days >50 µg m^{-3}). While these

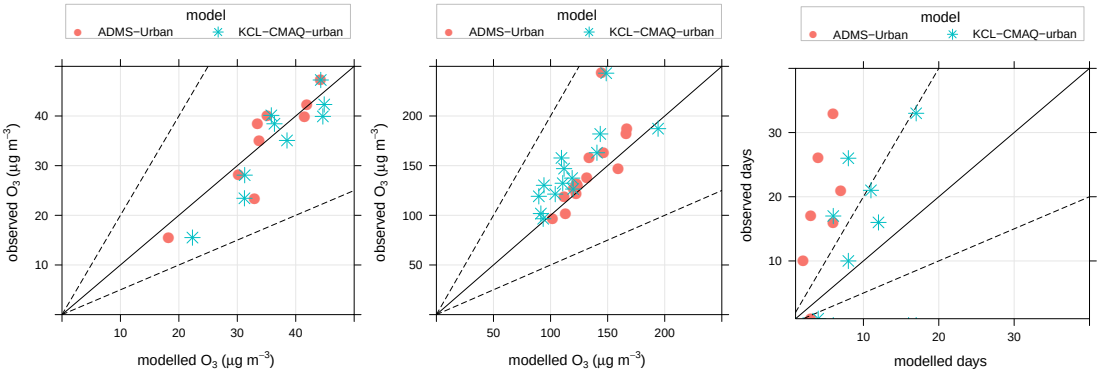

FIGURE 5.2: Comparison of modelled and observed annual mean, maximum daily 8-hour mean and number of days the rolling 8 hour mean is >100 µg m^{-3} for O_3 concentrations for ADMS-Urban and KCL CMAQ-Urban.

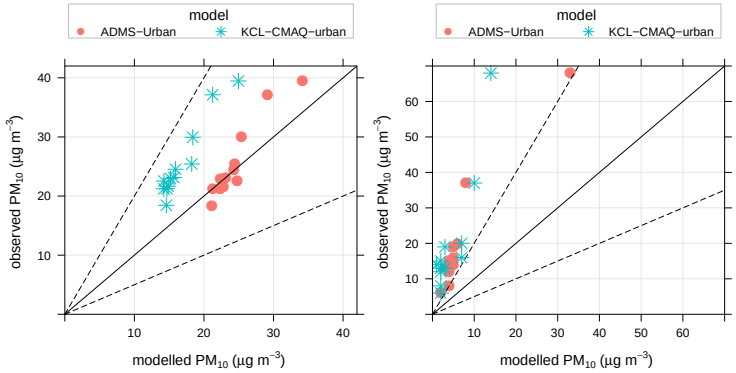

FIGURE 5.3: Comparison of modelled and observed annual mean and the number of days >50 µg m^{-3} for PM_{10} concentrations for ADMS-Urban and KCL CMAQ-Urban.

statistics are relevant from a health-evidence perspective they do not lend themselves to accurate prediction by models. It is the case however that these types of model *can* predict high concentration metrics well but their performance for threshold-type statistics tend to be much worse. For example, in Figure 5.2 it can be seen that both ADMS-Urban and KCL CMAQ-Urban predict the maximum daily 8-hour mean O_3 concentration well but are much worse at predicting the number of days where the threshold of 100 µg m^{-3} is exceeded.

5.3 A closer look at hourly predictions

Bivariate polar plots provide a useful approach for comparing models with measurements when hourly data are available (D. C. Carslaw, Beevers et al. 2006; D. C. Carslaw and Ropkins 2012). They provide a tough test of models because models not only need to estimate the absolute magnitudes of concentrations well, but they also need to capture the wind speed and wind direction characteristics of the dispersion. The latter point can provide information on whether sources have been modelled adequately or indeed whether sources even exist in the model to begin with. Three examples are given for NO_x that highlights these issues.

In Figure 5.4 polar plots are shown for the North Kensington background site. Both models do well in capturing the nature of this site in that the highest concentrations are predicted at the lowest wind speeds. This behaviour is typical of a background monitoring

sites where concentrations are strongly influenced by well-mixed sources and where a stable atmosphere (i.e. low wind speeds) leads to higher ground-level concentrations. There is a tendency of both models to overestimate NO_x concentrations at low wind speeds when the wind is from the east i.e. is from the direction of central London. This behaviour could be caused by insufficient dispersion of area or volume sources under low wind speed conditions in the model.

The ADMS-Urban model does tend to predict higher concentrations at low wind speeds, which could be due perhaps to differences in source apportionment. However, as Table 5.6 shows there is not much difference between the source apportionment of ADMS-Urban and ERG-Toolkit at this site.[5] The differences may arise due to how the sources have been parameterised in the first place e.g. initial mixing depth of volume sources. In should be noted that both models overestimate NO_x concentrations at North Kensington (measured = 49 µg m^{-3}; ADMS-Urban = 73 µg m^{-3} and KCL CMAQ-Urban = 64 µg m^{-3}).

Note that in Figure 5.4 to Figure 5.10 use is made of the meteorological data used in each model i.e. surface measurements from Heathrow for ADMS-Urban and the WRF model for KCL CMAQ-Urban.

TABLE 5.6: Mean source contributions to NO_x concentrations by model and major source category for the North Kensignton site (µg m^{-3}).

variable	ADMS-Urban	BRUTAL	ERG-Toolkit	PCM
cars	12.1	13.3	16.3	0.0
LGVs	5.5	3.8	5.2	0.0
HGVs	7.9	6.9	5.8	0.0
Bus	5.4	10.3	7.4	0.0
roads	0.0	0.0	0.0	31.5
gas	16.2	21.9	16.3	23.4
other.tran	13.2	12.7	10.8	9.4
background	13.0	11.5	13.5	10.8
other	0.9	1.5	0.5	0.8

Figure 5.5 shows a different characteristic of model performance where NO_x concen-

[5] This analysis assumes that ERG-Toolkit and KCL CMAQ-Urban results in a similar source apportionment.

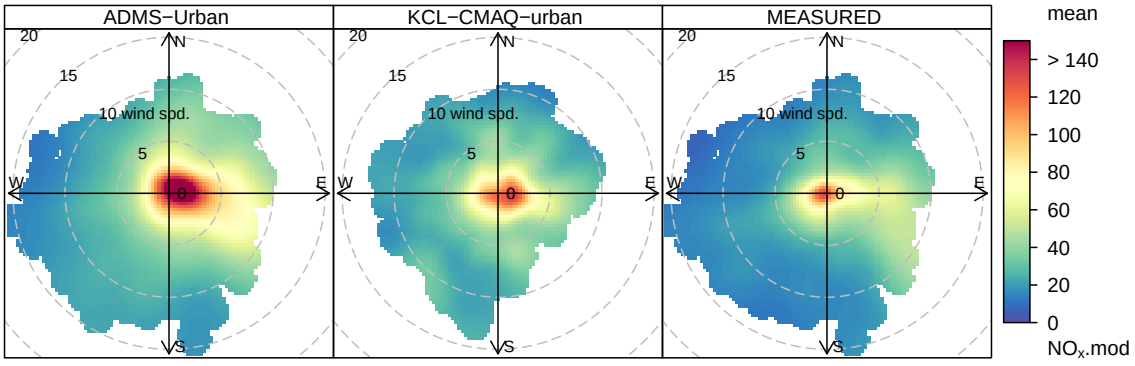

FIGURE 5.4: Bivariate polar plot of NO_x concentration at the North Kensington site comparing ADMS-Urban and the KCL CMAQ-Urban models with measurements.

trations at the London Bloomsbury site are considered. In this case both models fail to capture higher concentrations to the north-west, which other analysis has shown is potentially due to a local NO_x combustion source. Sources such as these are not adequately quantified in the LAEI and the poor model performance in this case is more to do with the emissions inventory rather than the models. Indeed, a consideration of other sites shows that there are many such sources that are inadequately quantified in the emissions inventories, which clearly limits the accuracy of model predictions.

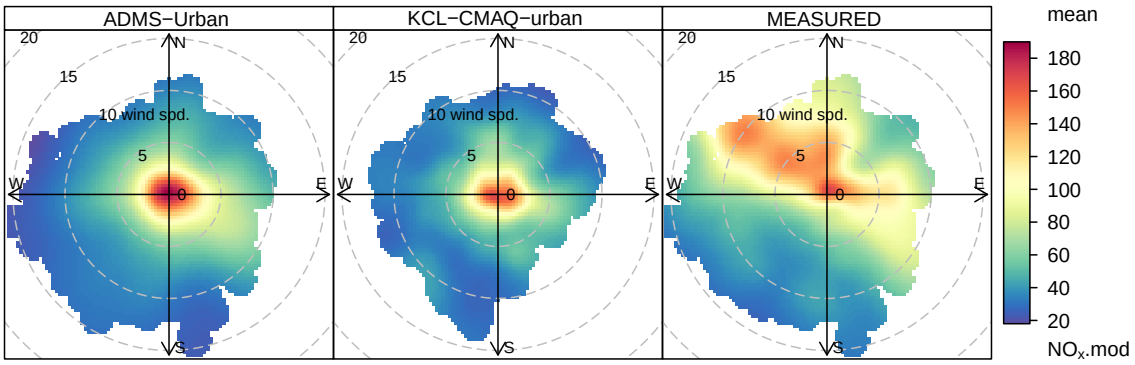

FIGURE 5.5: Bivariate polar plot of NO_x concentration at the London Bloomsbury site comparing ADMS-Urban and the KCL CMAQ-Urban models with measurements.

In the third example concentrations at the Marylebone Road site are considered. The Marylebone Road site is in a street canyon location and the configuration of the canyon has a strong effect on how vehicle emissions disperse. In particular, when the wind is from the south there is a strong re-circulation effect and concentrations remain high even at relatively high wind speeds, as shown on the right-hand panel of Figure 5.6. In this case, these complex mixing patterns are not captured by either model.[6] Clearly, it would not be expected that these models would capture very local influences due to complex dispersion and this again limits the performance of the models at sites where local dispersion is strongly influenced by local effects.

In fact, problems with the model predictions at this location can be seen in a scatter plot, as shown in Figure 5.7 for NO_x. Plotting the data in this way reveals that both models show a bimodal distribution where significant numbers of predictions are either above or below the measured value. Figure 5.7 indicates that at this location the models

[6]Note, both models effectively used the OSPM street canyon model in these situations; or variant of.

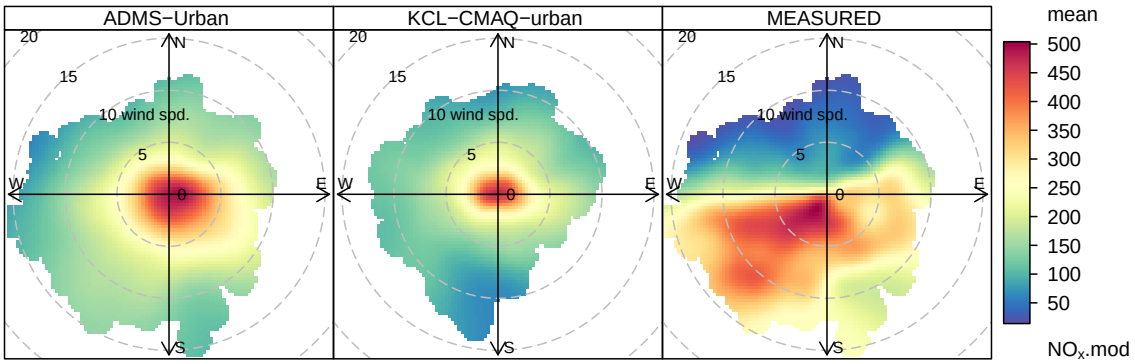

FIGURE 5.6: Bivariate polar plot of NO_x concentration at the Marylebone Road site comparing ADMS-Urban and the KCL CMAQ-Urban models with measurements.

tend to either over or under-predict concentrations. As it will be shown, this pattern of predictions is indicative of the models failing to capture the street canyon recirulation at this location.

While it is straightforward to compare models using various statistics and plots, it is harder to pin down the reasons for poor model performance. One method is to consider *conditional analysis*. Such an analysis can not only highlight where a model performs poorly — but importantly also considers the behaviour of other variables at the same time. By considering other variables in this way it can be easier to identify some of the potential reasons for poor model performance. In addition, analysing measurements in an identical way, so that they represent a 'perfect model' for comparison, it can be readily seen where model where models differ from reality.

Using these conditional approaches, the performance of the models at Marylebone Road can be analysed as shown in Figure 5.8. In Figure 5.8 the conditional quantile plot (left) shows that when the NO_x predictions are low, the observed NO_x is higher and when the predictions of NO_x are high the observed NO_x is lower. The plot is constructed by dividing the predicted concentrations into intervals. For each of these intervals the corresponding measurements are considered and their median values together with the spread in values (between the 25-75th and 10-90th percentile) are calculated. A perfect model would lie on the blue 1:1 line i.e. would have an identical median value and very small spread in concentrations. Indeed, a 'perfect' model is shown in the bottom panels of the plots in Figure 5.8. Taking predicted concentrations of about 700 µg m^{-3} both models show that the corresponding measurements tend to be lower (about 500 µg m^{-3}). The plot also shows for the interval around 700 µg m^{-3} there is a large spread in observed values — the 90th percentile range covers a range from 100 to 800 µg m^{-3}.

Also shown in Figure 5.8 is the *corresponding* proportion of wind speed intervals, which highlights that when NO_x is underestimated ($\lessapprox$200 µg m^{-3} NO_x) the wind speed tends to be high — and when NO_x is overestimated the wind speed tends to be low for both models. The performance of both models is very similar. The measured values (bottom plots of Figure 5.8) show that across most of the range of NO_x concentrations there is a roughly equal split in corresponding wind speed intervals. It is only for very high NO_x

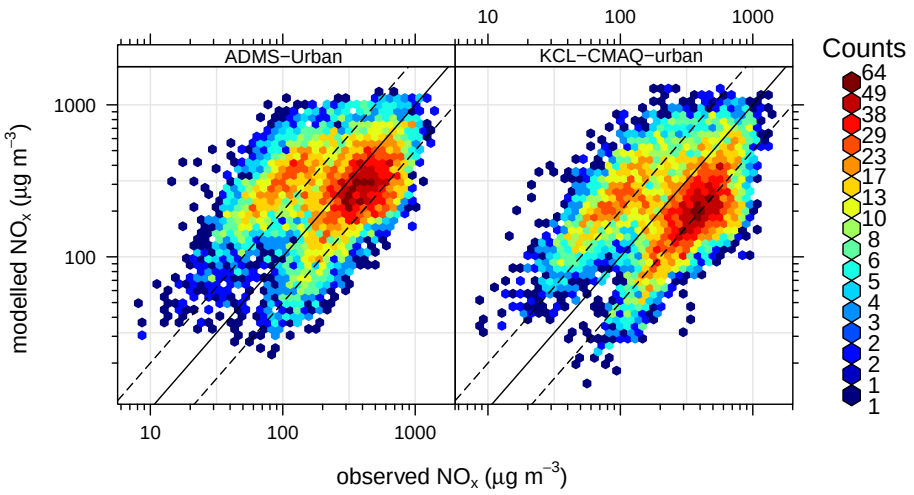

FIGURE 5.7: Scatter plot of modelled vs measured concentration of NO_x at Marylebone Road (µg m^{-3}). The plot shows the frequency of points and helps to reveal where the bulk of the predictions are located. The dashed lines show the factor of two envelope.

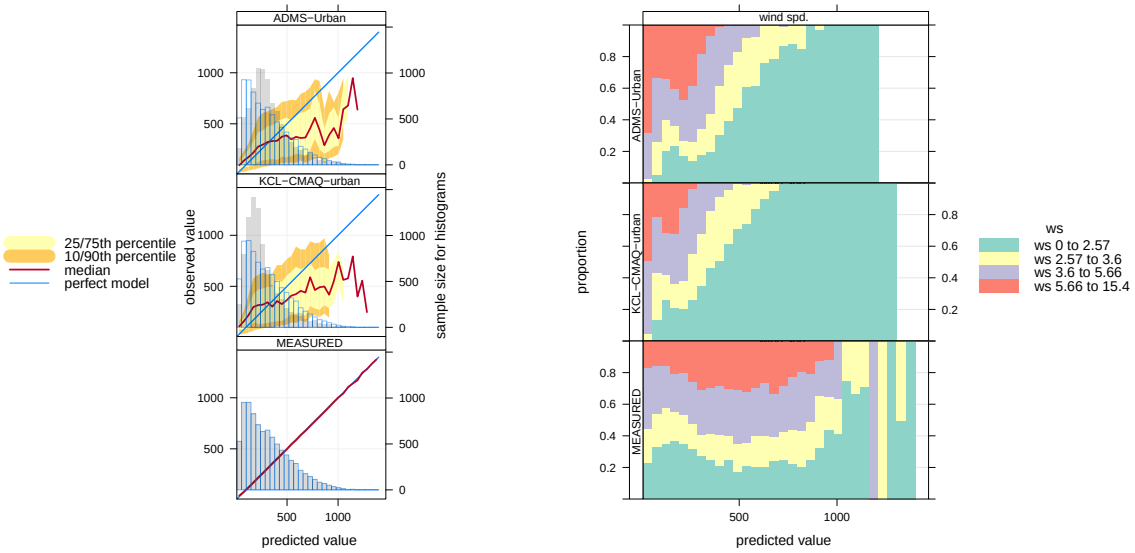

FIGURE 5.8: Left: conditional quantile plot of NO_x performance at Marylebone Road. The blue line shows the results for a perfect model. The red line shows the median value of the predictions. The shading shows the predicted quantile intervals i.e. the 25/75th and the 10/90th. A perfect model would lie on the blue line and have a very narrow spread. There is still some spread because even for a perfect model a specific quantile interval will contain a range of values. However, for the number of bins used in this plot the spread will be very narrow, as shown by the results in the "measured" panel. Finally, the shaded histogram shows the counts of predicted values and the blue histogram the counts of the observed values. Right: the plot shows the proportion of NO_x prediction intervals by wind speed interval.

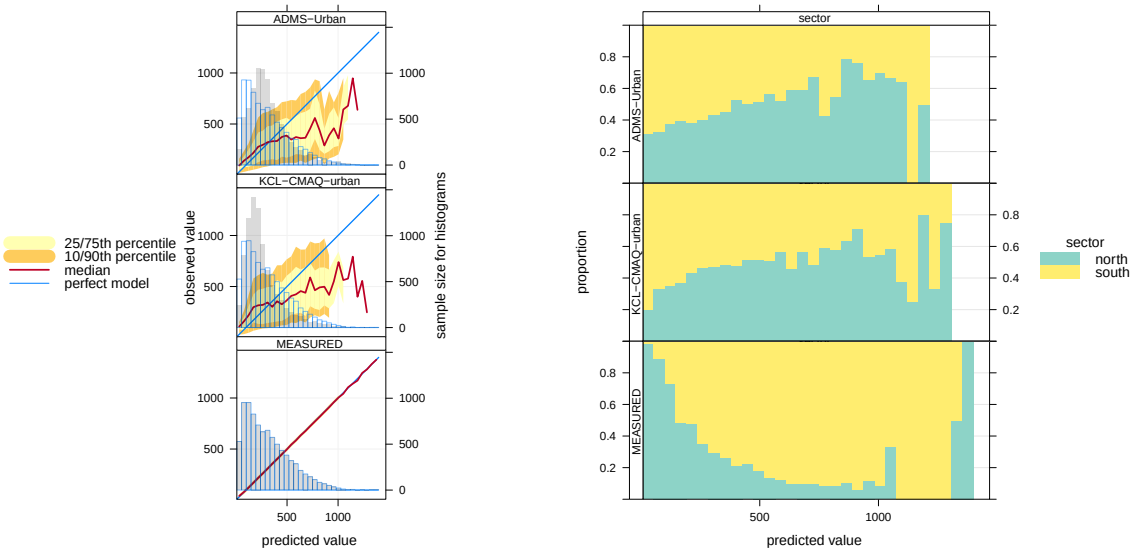

FIGURE 5.9: Left: conditional quantile plot of NO_x performance at Marylebone Road. The blue line shows the results for a perfect model. The red line shows the median value of the predictions. The shading shows the predicted quantile intervals i.e. the 25/75th and the 10/90th. A perfect model would lie on the blue line and have a very narrow spread. There is still some spread because even for a perfect model a specific quantile interval will contain a range of values, as shown by the results in the "measured" panel. However, for the number of bins used in this plot the spread will be very narrow. Finally, the shaded histogram shows the counts of predicted values and the blue histogram the counts of the observed values. Right: the plot shows the proportion of NO_x prediction intervals by wind sector.

concentrations that low wind speeds are important for measured values. The most striking feature of the measurements (but not the predictions) is that when NO_x concentrations are high (>500 μg m^{-3}) there is also a relatively high proportion of high wind speeds. For example, when the concentration of NO_x is ≈500 μg m^{-3} approximately 25% of the time the wind speed is in the highest interval (5.6–15.5 m s^{-1}) — but there are no such conditions present for predicted concentrations of NO_x at this level.

The conditional quantile plots are shown for wind sector (northerly and southerly component wind direction) in Figure 5.9. This Figure shows for measured concentrations (bottom plots of Figure 5.9), when the concentration of NO_x is low the wind direction is dominated by northerly winds. This behaviour is in contrast to the predictions from both models, where low NO_x concentration predictions tend to be associated with a higher proportion of southerly winds. Conversely, as concentrations of NO_x increase there is an increasing proportion of southerly winds.

Figure 5.8 and Figure 5.9 highlight that the models fail to adequately capture the strong street canyon recirculation at Marylebone Road. When the wind is from the south the recirculation results in high concentrations of NO_x as the wind blows the vehicle emissions toward the monitor. Conversely, when the wind is from the north, 'fresh' air is brought down to the monitor resulting in lower NO_x concentrations. It should be noted that at most sites both models tend to capture the wind speed and direction characteristics of measured concentrations quite well. Nevertheless, sites such as Marylebone Road are important because concentrations and exceedances of EU Limit Values are higher than at many other sites and therefore it is important that models are capable of producing reasonable predictions.

Both models actually predict annual mean NO_x at Marylebone Road well (measured = 312 μg m^{-3}; ADMS-Urban = 321 μg m^{-3} and KCL CMAQ-Urban = 266 μg m^{-3}). However, these results hide the fact that the predictions are arrived at by overestimating concentrations considerably when the wind is from the north and underestimating when the wind is from the south. These characteristics will also affect estimates of other statistics such as the number of hours NO_2 is >200 μg m^{-3}, for example. Indeed it can be shown that based on measurements 95% of the hours where NO_2 is >200 μg m^{-3} the wind is from the south. For the models, ADMS-Urban and KCL CMAQ-Urban the proportion is 46% i.e. roughly half the hours when NO_2 >200 μg m^{-3} occur when the wind in from the north. There are many situations in modelling such as this i.e. where a model 'gets it right for the wrong reasons'. The extent to which this matters will depend on the context and it may be reasonable to accept such models represent 'typical' conditions. For example, it is likely that if the Marylebone Road site was located on the north side of the road, exceedances of the hourly mean NO_2 Limit Value would be very low — but the models would predict a much higher number. These issues are however important for some situations when attempting to account for model performance at specific receptor locations.

Figure 5.10 shows the Conditional Probability Function (CPF) for NO_2 concentrations, with the aim of highlighting those conditions when the hourly limit value for NO_2 is exceeded. This plot shows the probability of NO_2 concentrations exceeding the 94th percentile, which in this case has an equivalent NO_2 concentration of 201 μg m^{-3} i.e. very close to the hourly limit value. What is clear from Figure 5.10 is that for measured concentrations the highest probabilities are associated with winds from the south and in particular the south-west. The probability of there being concentrations greater than 201 μg m^{-3} NO_2 when the wind is from the north are effectively zero; which

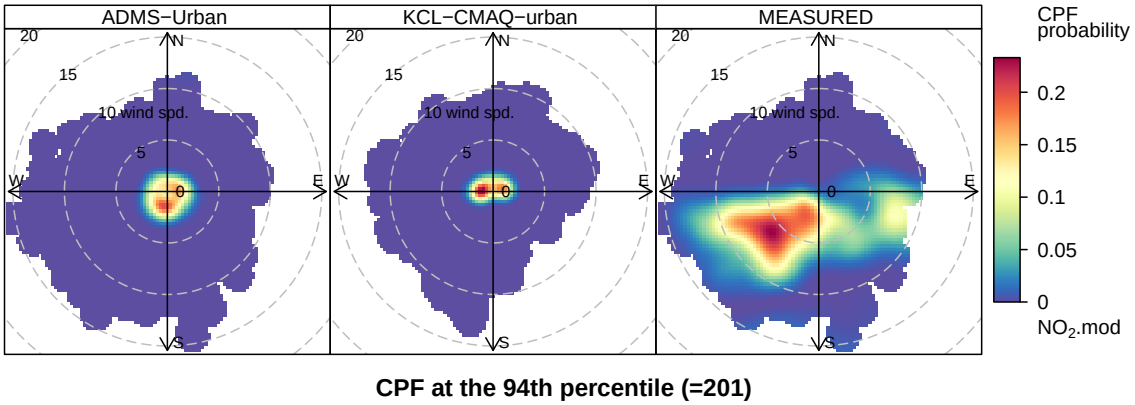

FIGURE 5.10: Conditional Probability Function applied to a bivariate polar plot of NO_2 concentration at the Marylebone Road site comparing ADMS-Urban and the KCL CMAQ-Urban models with measurements. The 94th percentile is considered, which is equivalent to hourly concentrations of NO_2 >200 µg m^{-3}.

again emphasises the importance of street canyon recirculation on NO_2 exceedances. Furthermore, these high concentrations are possible even under strong wind speeds up to about 10 m s^{-1}. In contrast the models show that concentrations of NO_2 only exceed 201 µg m^{-3} under very low wind speed conditions (less than ≈2 m s^{-1}); from almost any wind direction. In other words, the conditions leading to exceedances of the hourly limit value for NO_2 differ markedly between the measurements and models.

5.3.1 OSPM model predictions

Both the ADMS-Urban and KCL CMAQ-Urban models use an implementation of the OSPM street canyon model (Berkowicz 2000). OSPM has been extensively tested and used throughout Europe (Kakosimos et al. 2010; Kukkonen, Partanen et al. 2003; Kukkonen, Valkonen et al. 2000). These validation studies generally show that OSPM performs well in a wide range of situations. It is useful to know therefore how well the *direct* use of the model captures the joint wind speed-direction variation of concentrations in a street canyon location. It is beyond the scope of this report to set up OSPM for a specific location such as Marylebone Road. However, it is instructive to apply the same techniques as used in this report to OSPM results from a street canyon location in Denmark. Data were obtained from Aarhus University, Denmark for 2009.[7]

Figure 5.12 shows the polar plot for the Jagtvej site in Copenhagen. The site is located on a straight road at a +30 degrees orientation to north on the east side of the road. Street canyon recirculation is apparent in both the measured and modelled concentration distribution shown in Figure 5.12 i.e. the highest concentrations of NO_x when the wind is from the east. This concentration distribution is in marked contrast to that seen in Figure 5.6.

While the Jagtvej site might be considered as 'ideal' i.e. a uniform street canyon where the OSPM model has been extensively tested, it does nevertheless capture the important recirculation that neither ADMS-Urban nor KCL CMAQ-Urban do. Clearly, further investigations of the use of OSPM or similar canyon models for use in the UK would be beneficial.

[7]The data can be downloaded from http://www.dmu.dk/en/air/models/background/ospmtool/.

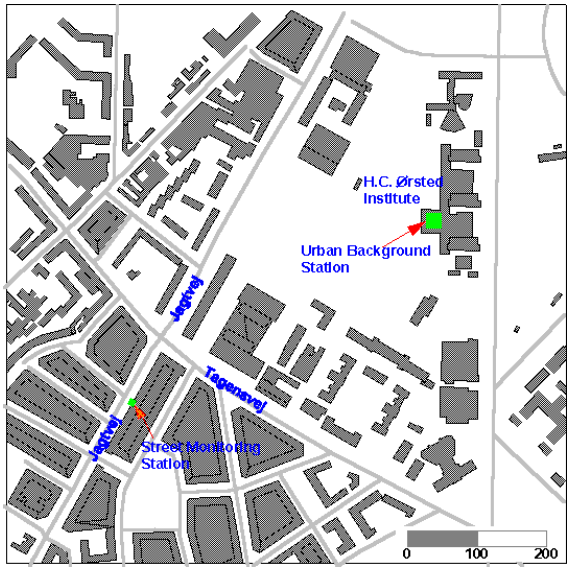

FIGURE 5.11: Map showing the location of the Jagtvej air pollution monitoring site in Copenhagen.

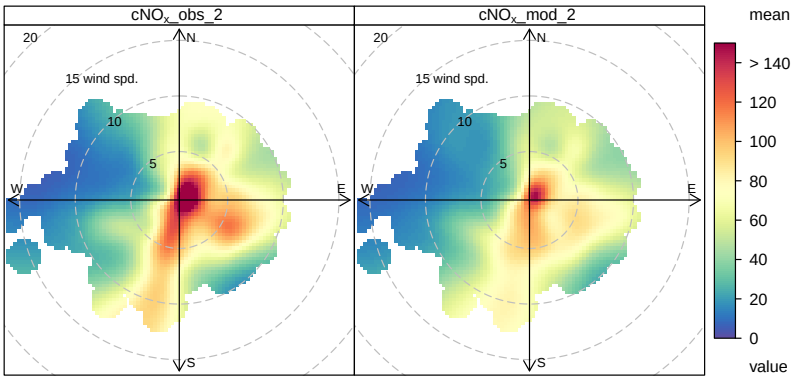

FIGURE 5.12: Comparison of measured and modelled NO_x using the OSPM model.

5.4 Hourly NO_x–NO_2 relationships

The hourly models allow an assessment to be made of the relationship between NO_x and NO_2. For each site the concentration of NO_x has been divided into intervals (0–5, 5–10 µg m^{-3}, ...) and the mean value of NO_2 calculated. The ratio of NO_2:NO_x was then calculated by site and NO_x interval. Table 5.7 gives the model evaluation statistics for the NO_2:NO_x ratio, split by site type and model. Overall, the KCL CMAQ-Urban model tends to have better performance, but the differences are small for most site types. At kerbside sites, KCL CMAQ-Urban tends to underestimate the NO_2/NO_x ratio where the *NMB* is −0.09. However, the negative bias for ADMS-Urban is larger at −0.20. At roadside sites KCL CMAQ-Urban has a small positive bias and ADMS-Urban a small negative bias. For other site types there is more agreement between the two models.

TABLE 5.7: Performance of ADMS-Urban and KCL CMAQ-Urban for NO_2:NO_x ratio by site type and model. Note for site type Krbs = kerbside, Rdsd = roadside, Sbrb = suburban and UrbB = urban background.

site type	model	n	FAC2	MB	MGE	NMB	NMGE	RMSE	r	COE
Krbs	KCL-CMAQ-urban	426	1.00	-0.04	0.05	-0.09	0.12	0.06	0.91	0.40
Krbs	ADMS-Urban	457	1.00	-0.08	0.08	-0.20	0.20	0.10	0.85	-0.08
Rdsd	KCL-CMAQ-urban	714	1.00	0.02	0.04	0.05	0.09	0.05	0.97	0.67
Rdsd	ADMS-Urban	839	1.00	-0.01	0.04	-0.03	0.11	0.06	0.93	0.64
Sbrb	KCL-CMAQ-urban	205	1.00	0.04	0.05	0.07	0.09	0.06	0.98	0.72
Sbrb	ADMS-Urban	318	0.99	0.03	0.06	0.07	0.13	0.07	0.96	0.68
UrbB	KCL-CMAQ-urban	293	1.00	0.00	0.03	0.01	0.06	0.04	0.98	0.82
UrbB	ADMS-Urban	381	1.00	-0.02	0.05	-0.04	0.11	0.06	0.98	0.72

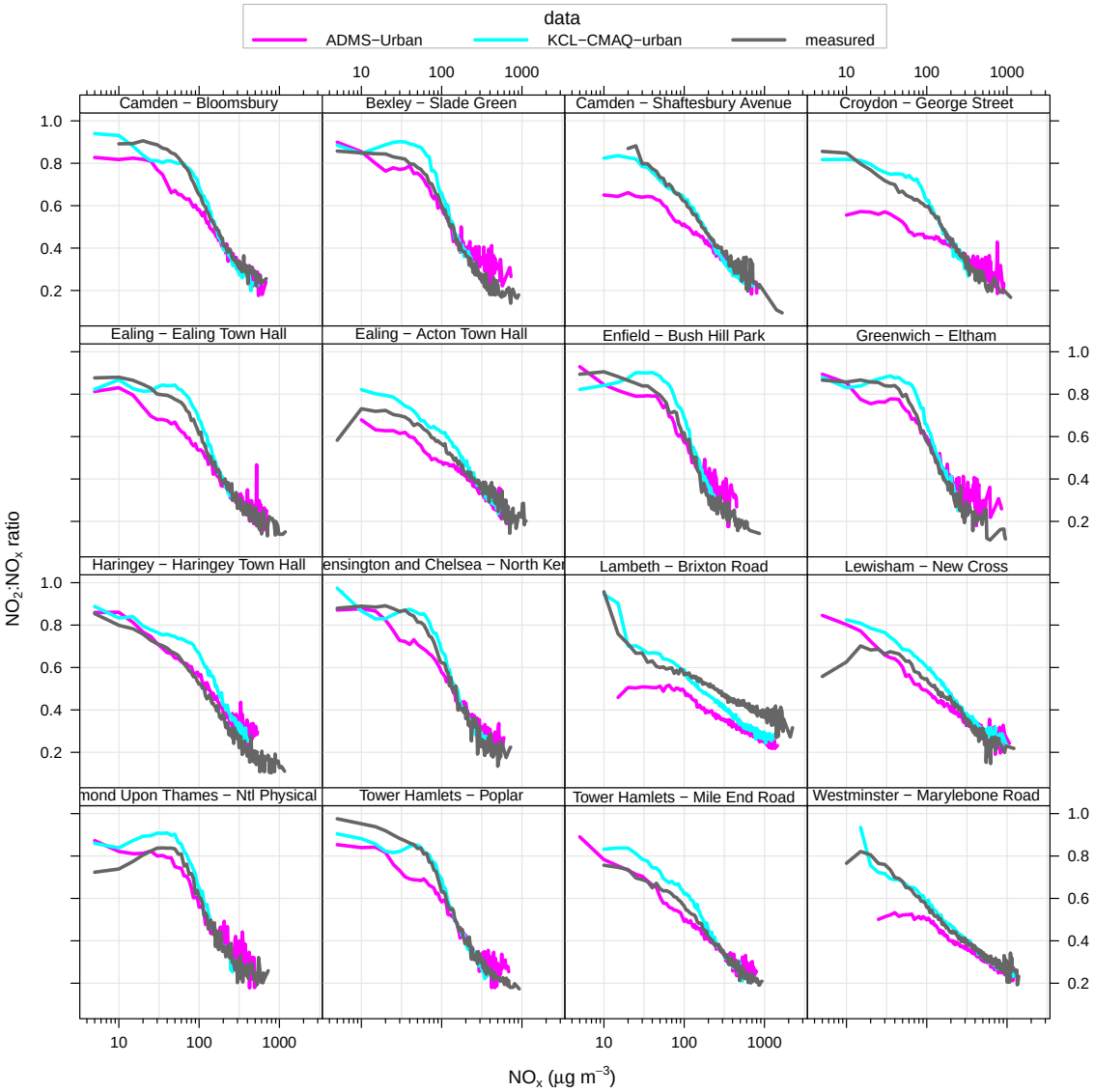

FIGURE 5.13: Relationship between NO_2/NO_x ratio and NO_x concentration for ADMS-Urban, KCL CMAQ-Urban and measurements. Note that the x-axis is on a $\log_{10}$ scale.

A summary of the relationship between NO_x and NO_2/NO_x for all sites is shown in Figure 5.13. The results show that both models tend to capture the NO_2/NO_x ratio well at each site, but there is a tendency for ADMS-Urban to underestimate the ratio. At higher concentrations of NO_x where primary NO_2 will be more important (particularly at

roadside and kerbside sites) there tends to be better agreement between the two models, suggesting the difference in approach to NO_2-NO-O_3 chemistry has an important effect controlling the differences between the models.

Taken as a whole however, both ADMS-Urban and KCL CMAQ-Urban capture the important features of the relationship between NO_x and NO_2 across a wide range of site types and NO_x and NO_2 concentrations. Some of the differences will also be due to assumptions regarding the primary NO_2 emission. The ADMS-Urban model for road vehicles assumes a fixed value of 22% for primary NO_2 for all road types. KCL CMAQ-Urban varies the primary NO_2 assumed by vehicle type and therefore treats each road specifically. These different approaches could account for some of the differences seen in the comparisons of NO_2/NO_x ratio. An advantage both these models have over models that predict annual means is the ability to consider the relationships between variables in more depth. The good performance in describing the non-linear relationship between NO_x and NO_2 concentration across a wide range of sites (and wide range of atmospheric conditions from NO_x to O_3-limited), also provides confidence that the changes predicted by these models are also likely to be good.

5.5 Performance against the new Daily Air Quality Index

This section briefly considers the performance of the models against the Defra new daily air quality index. More details concerning the index can be found at `http://uk-air.defra.gov.uk/air-pollution/daqi` (COMEAP 2011). The air quality index is shown in Table 5.8. The index is more relevant to air quality forecasting, but both models considered in this section can in principle be used (or are used) for this purpose and it is worth considering their performance. Note that the indexes are calculated for different averaging times dependent on the pollutant: rolling 8-hour mean for O_3, hourly means for NO_2 and a fixed 24-hour mean for PM_{10} and $PM_{2.5}$.

A common approach to evaluate the quality of forecast predictions is to consider the categorical prediction e.g. was a 'high' prediction forecast by the model? In this case the response is yes or no. There are various methods for evaluating the quality of predictions for these cases and Stephenson (2000) explores approaches based on the 'odds ratio'. Specifically, the *odds ratio skill score* is used as a measure of the predictive performance of models for these situations i.e. where there is a yes/no, TRUE/FALSE outcome. As mentioned above, these issues are more relevant to models used for forecasting where the policy maker would like to know whether to issue a warning that a certain level of concentration is likely to be exceeded. This section only considers whether the models show much bias when it comes to predicting the daily air quality indexes, but could be extended to consider the odds ratio skill score if it was considered important and useful.

Figure 5.14 to Figure 5.17 show the mean bias (MB) and normalised mean bias (NMB) for NO_2, O_3, PM_{10} and $PM_{2.5}$ respectively. In these plots the 'extreme' Lambeth 4 site has been removed from the comparison. In general it can be seen that the lowest concentrations tend to be over-estimated by both models most of the time (index = 1), but the overall tendency is that both models tend to show a negative bias as concentrations of all species increase. For the most part the performance of the two models is similar, but the ADMS-Urban model does show less bias for PM_{10}. Again, this might be expected because CMAQ estimates of PM_{10} concentration are generally underestimated.

The implications of these results are that the models typically tend to underestimate the daily air quality index by 2 to 4 levels. A few points should be noted however. First, the

TABLE 5.8: The new Defra daily air quality index (values in μg m^{-3}).

Band	Description	NO_2	O_3	PM_{10}	$PM_{2.5}$
1	Low	0–66	0–33	0–16	0–11
2	Low	67–133	34–65	17–33	12–23
3	Low	134–199	66–99	34–49	24–34
4	Moderate	200–267	100–120	50–58	35–41
5	Moderate	268–334	121–140	59–66	42–46
6	Moderate	335–399	141–159	67–74	47–52
7	High	400–467	160–187	75–83	53–58
8	High	468–534	188–213	84–91	59–64
9	High	535–599	214–239	92–99	65–69
10	Very High	600 or more	240 or more	100 or more	70 or more

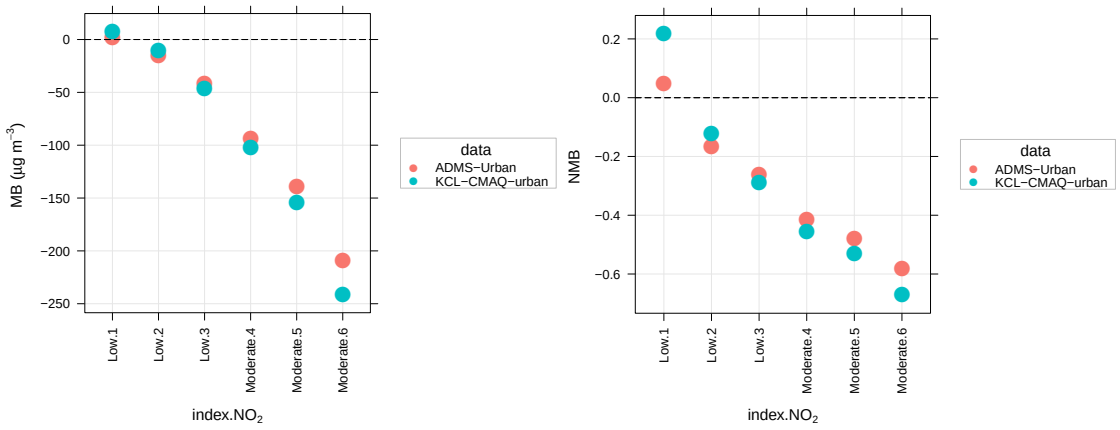

FIGURE 5.14: Comparison of model performance (mean bias, *MB* and normalised mean bias, *NMB*) against the new air quality index for NO_2 concentrations.

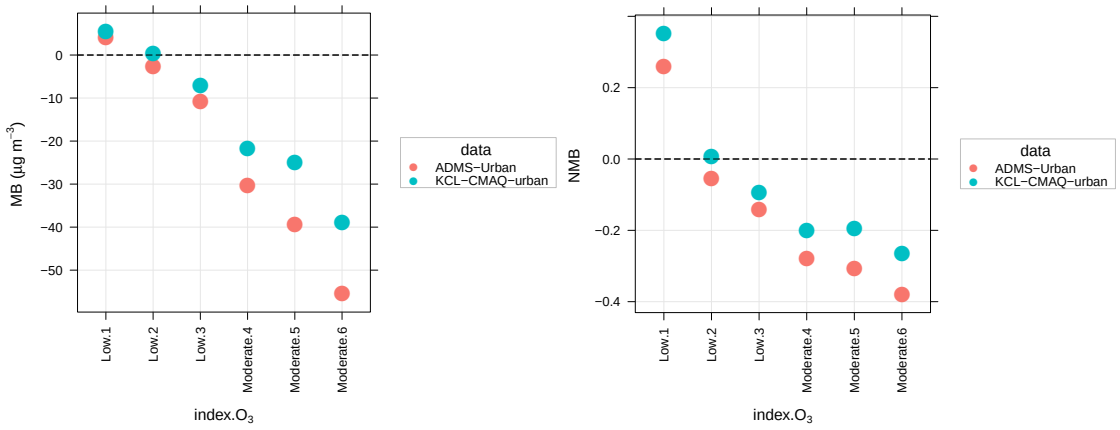

FIGURE 5.15: Comparison of model performance (mean bias, *MB* and normalised mean bias, *NMB*) against the new air quality index for O_3 concentrations.

analysis in this section is specific to the 15 receptor locations considered in the analysis. Second, if the models were actually used in 'forecast mode' they would rely on forecast meteorology rather than local surface measurements. The uncertainty in the predictions could therefore be higher than is given here. It should also be noted that these comparisons require the models to get the *timing* of the predicted index correct.

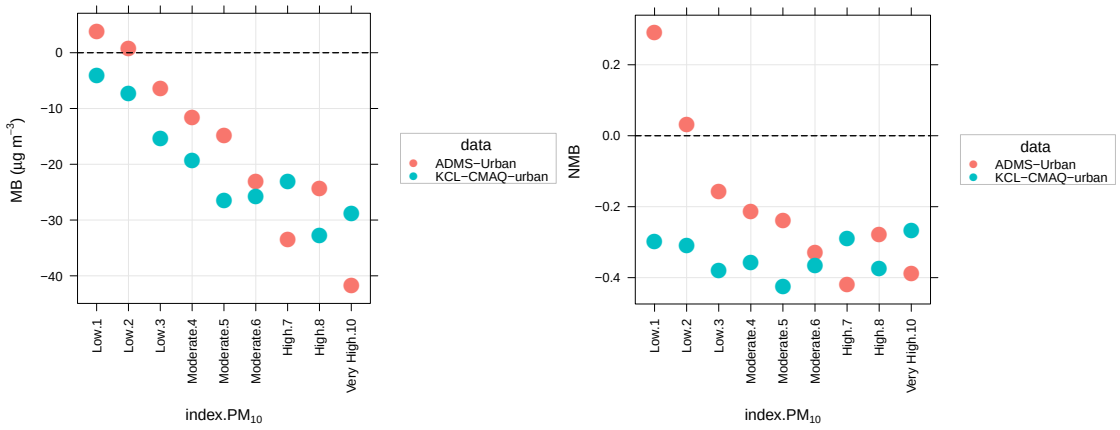

FIGURE 5.16: Comparison of model performance (mean bias, *MB* and normalised mean bias, *NMB*) against the new air quality index for PM_{10} concentrations.

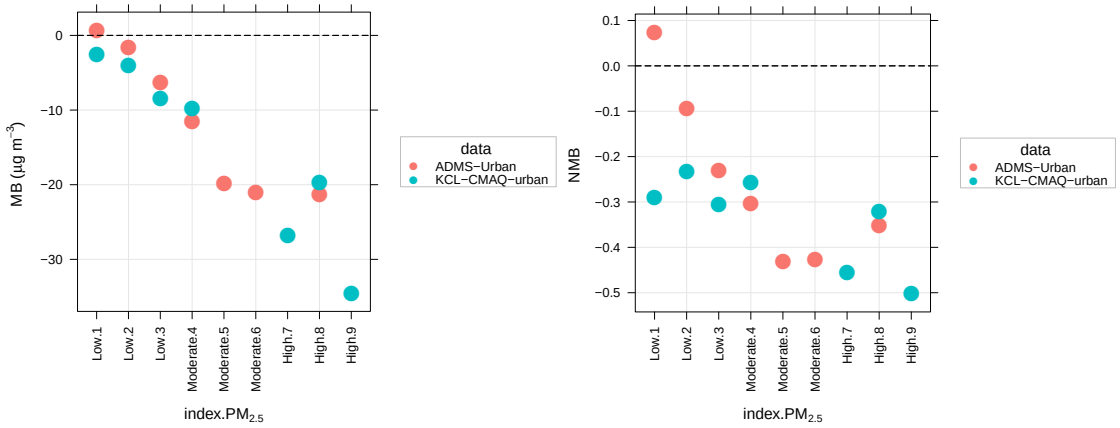

FIGURE 5.17: Comparison of model performance (mean bias, *MB* and normalised mean bias, *NMB*) against the new air quality index for $PM_{2.5}$ concentrations.

The forecasting skill of ADMS-Urban and other models will be considered in more depth in subsequent work.

Acknowledgements

The assistance provided by Dr Matthias Ketzel from Aarhus University, Denmark on the OSPM model is gratefully acknowledged.

References

AKIMA, H. (2012). *akima: Interpolation of irregularly spaced data. Fortran code by H. Akima R port by Albrecht Gebhardt aspline function by Thomas Petzoldt <thomas.petzoldt@tu-dresden.de> enhancements and corrections by Martin Maechler*. R package version 0.5-7 (cit. on p. 21).

AQEG (2012). *Fine Particulate Matter ($PM_{2.5}$) in the United Kingdom*. Air Quality Expert Group. Report prepared by the Air Quality Expert Group for the Department for Environment, Food, Rural Affairs; Scottish Executive; Welsh Assembly Government; and Department of the Environment in Northern Ireland. (cit. on p. 9).

Beevers, S. D., N. Kitwiroon, M. L. Williams and D. C. Carslaw (2012). "One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions". In: *Atmospheric Environment* 59, pp. 47–58. doi: 10.1016/j.atmosenv.2012.05.034 (cit. on p. 11).

Berkowicz, R. (2000). "OSPM — A Parameterised Street Pollution Model". In: *Environmental Monitoring and Assessment* 65.1-2, pp. 323–331 (cit. on p. 43).

Carslaw, D. C. and S. D. Beevers (2005). "Estimations of road vehicle primary NO_2 exhaust emission fractions using monitoring data in London". In: *Atmospheric Environment* 39.1, pp. 167–177 (cit. on p. 11).

Carslaw, D. C., S. D. Beevers, K. Ropkins and M. C. Bell (2006). "Detecting and quantifying aircraft and other on-airport contributions to ambient nitrogen oxides in the vicinity of a large international airport". In: *Atmospheric Environment* 40.28, pp. 5424–5434 (cit. on p. 37).

Carslaw, D. C. and K. Ropkins (2012). "openair — An R package for air quality data analysis". In: *Environmental Modelling & Software* 27–28, pp. 52–61. doi: 10.1016/j.envsoft.2011.09.008 (cit. on pp. 9, 37).

Carslaw, D., P. Agnew, S. Beevers, C. Chemel, S. Cooke, L. Davis, D. Derwent, X. Francis, A. Fraser, N. Kitwiroon, T. Murrells, A. Redington, N. Savage, R. Sokhi and M. Vieno (2013). *Defra Phase 2 Regional Model Evaluation*. King's College London. (cit. on p. 34).

COMEAP (2011). *Review of the UK Air Quality Index: A report by the Committee on the Medical Effects of Air Pollutants* (cit. on p. 46).

Derwent, D., A. Fraser, J. Abbott, M. Jenkin, P. Willis and T. Murrells (2010). *Evaluating the Performance of Air Quality Models*. Issue 3/June 2010 (cit. on p. 53).

Grice, S. E., J. N. Lingard, J. R. Stedman, S. L. Cooke, Yap, F. W., A. J. Kent, T. J. Bush, K. J. Vincent and J. Abbott (2010). *UK air quality modelling for annual reporting 2008 on ambient air quality assessment under Council Directives 96/62/EC, 1999/30/EC and 2000/69/EC. Report to The Department for Environment, Food and Rural Affairs, Welsh Assembly Government, the Scottish Executive and the Department of the Environment for Northern Ireland. AEA report. AEAT/ENV/R/2656 Issue 1.* http://uk-air.defra.gov.uk/reports/cat09/1007201636_dd122008mapsrep_v4.pdf (cit. on p. 12).

Jenkin, M. E. (2004a). "Analysis of sources and partitioning of oxidant in the UK–Part 1: the NO_x-dependence of annual mean concentrations of nitrogen dioxide and ozone". In: *Atmospheric Environment* 38.30, pp. 5117–5129 (cit. on p. 12).

– (2004b). "Analysis of sources and partitioning of oxidant in the UK–Part 2: contributions of nitrogen dioxide emissions and background ozone at a kerbside location in London". In: *Atmospheric Environment* 38.30, pp. 5131–5138 (cit. on p. 12).

Kakosimos, K., O. Hertel, M. Ketzel and R. Berkowicz (2010). "Operational Street Pollution Model (OSPM)–a review of performed application and validation studies, and future prospects". In: *Environmental Chemistry* 7.6, pp. 485–503 (cit. on p. 43).

Kelly, F., H. R. Anderson, B. Armstrong, R. Atkinson, B. Barratt, S. Beevers, D. Derwent, D. Green, I. Mudway and P. Wilkinson (2011). *Part 1. Emissions modeling and analysis of air pollution measurements. In: The Impact of the Congestion Charging Scheme on Air Quality in London. Research Report 155. Health Effects Institute, Boston, MA*. (Cit. on p. 11).

Kukkonen, J., L. Partanen, A. Karppinen, J. Walden, R. Kartastenpaa, P. Aarnio, T. Koskentalo and R. Berkowicz (2003). "Evaluation of the OSPM model combined

with an urban background model against the data measured in 1997 in Runeberg Street, Helsinki". In: *Atmospheric Environment* 37.8, pp. 1101–1112 (cit. on p. 43).

Kukkonen, J., E. Valkonen, J. Walden, T. Koskentalo, P. Aarnio, A. Karppinen, R. Berkowicz and R. Kartastenpaa (2000). "A measurement campaign in a street canyon in Helsinki and comparison of results with predictions of the OSPM model". In: *Atmospheric Environment* 35.2, pp. 231–243 (cit. on p. 43).

Legates, D. R. and G. J. McCabe Jr (1999). "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation". In: *Water Resources Research* 35.1, pp. 233–241 (cit. on pp. 10, 54).

Legates, D. R. and G. J. McCabe (2012). "A refined index of model performance: a rejoinder". In: *International Journal of Climatology* (cit. on pp. 10, 54).

McHugh, C. A., D. J. Carruthers and H. A. Edmunds (1997a). "ADMS and ADMS-Urban". In: *International Journal of Environment and Pollution* 8.3-6, pp. 438–440 (cit. on p. 10).

– (1997b). "ADMS-Urban: an air quality management system for traffic, domestic and industrial pollution". In: *International Journal of Environment and Pollution* 8.3-6, pp. 666–674 (cit. on p. 10).

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria (cit. on p. 9).

Stephenson, D. B. (2000). "Use of the "odds ratio" for diagnosing forecast skill". In: *Weather and Forecasting* 15.2, 221–232 (cit. on p. 46).

Xie, Y. (2013a). *Dynamic Documents with R and knitr*. ISBN 978-1482203530. Chapman and Hall/CRC (cit. on p. 9).

– (2013b). *knitr: A general-purpose package for dynamic report generation in R*. R package version 1.1 (cit. on p. 9).

A Source apportionment plots

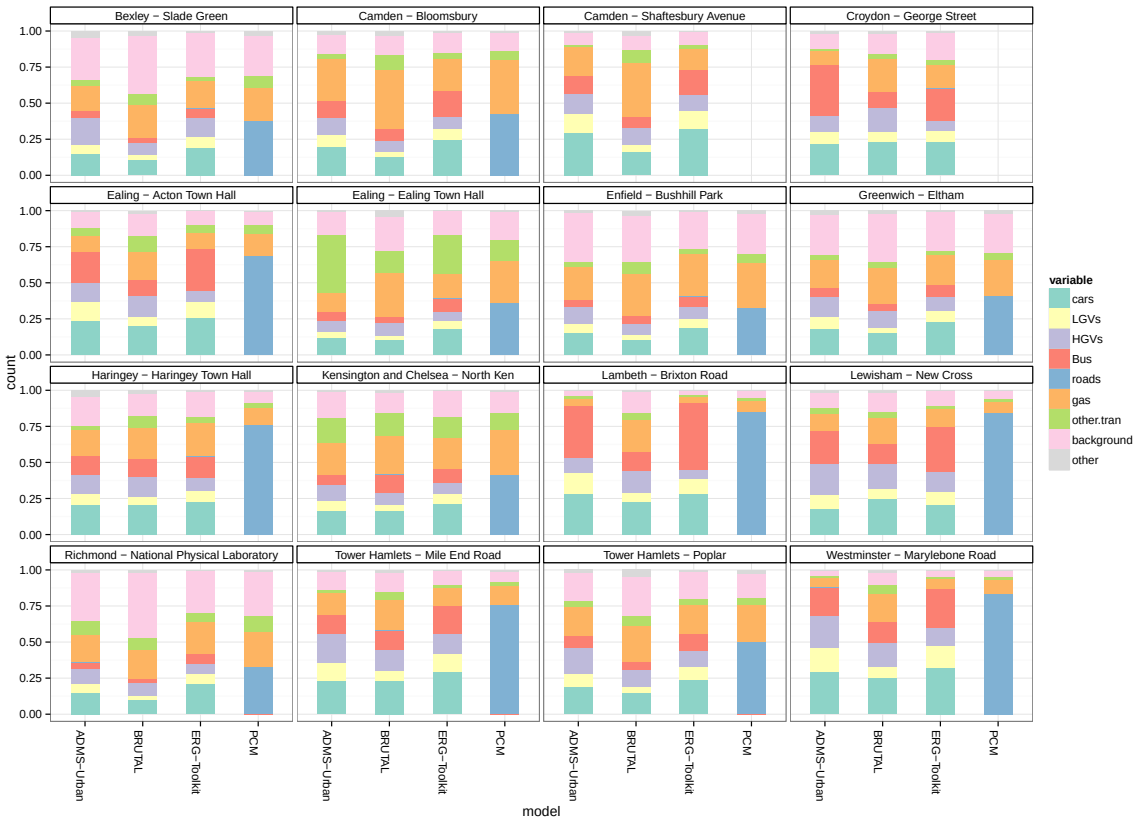

FIGURE A.1: Source apportionment of NO_x concentrations at 16 sites scaled to equal unity to emphasise relative contributions.

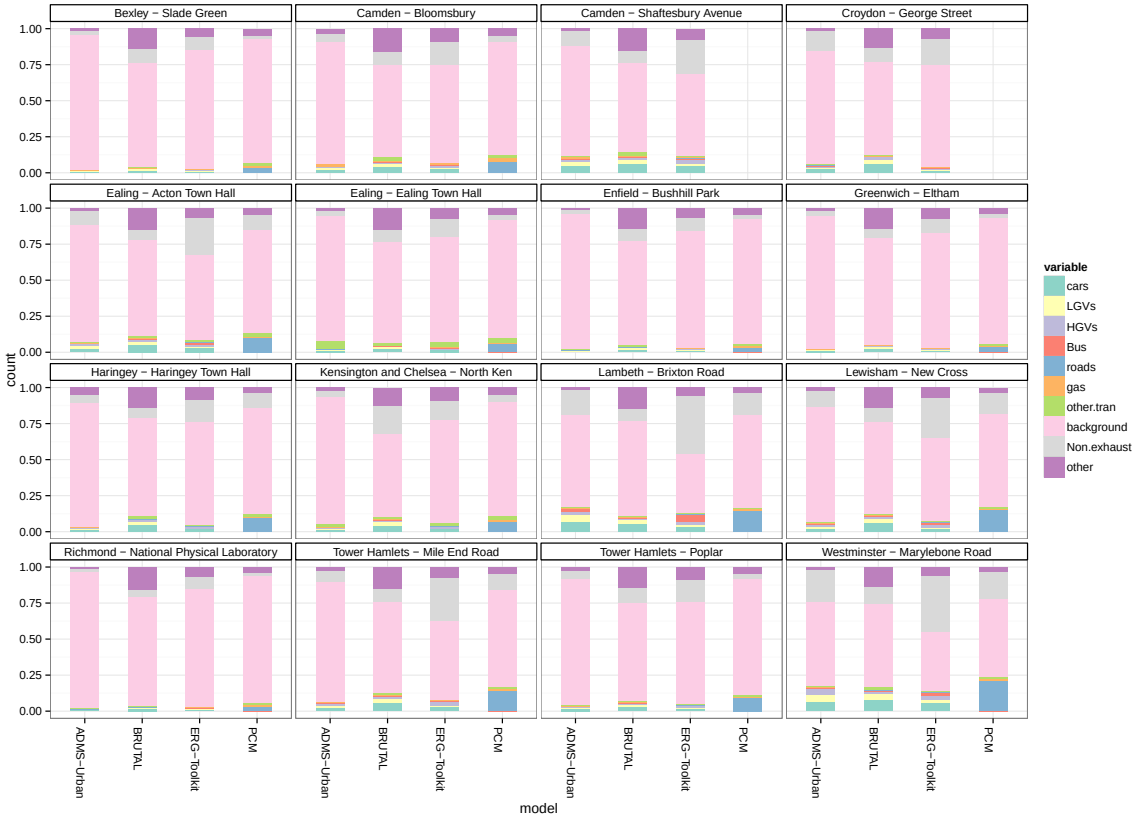

FIGURE A.2: Source apportionment of PM_{10} concentrations at 16 sites scaled to equal unity to emphasise relative contributions.

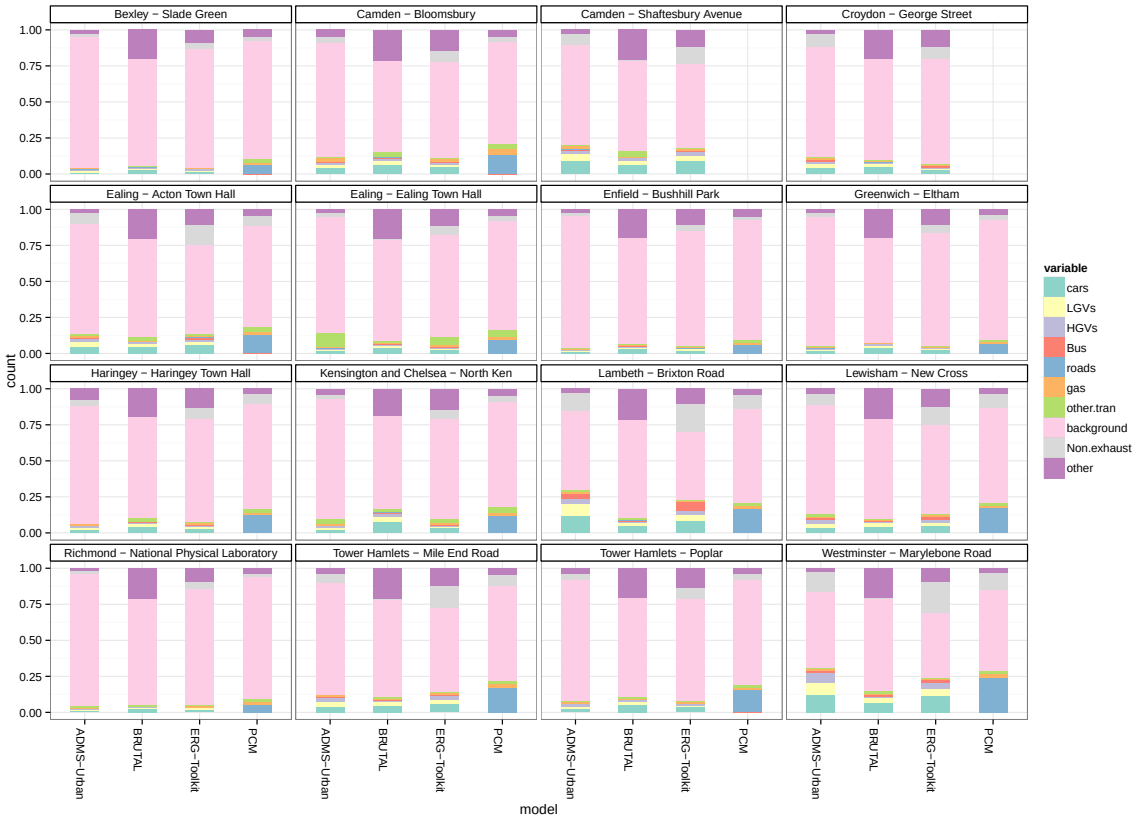

FIGURE A.3: Source apportionment of $PM_{2.5}$ concentrations at 16 sites scaled to equal unity to emphasise relative contributions.

B Model performance evaluation statistics

There are a very wide range of evaluation statistics that can be used to assess model performance. There is, however, no single statistic that encapsulates all aspects of interest. For this reason it is useful to consider several performance statistics and also to understand the sort of information or insight they might provide. The performance statistics used here have mostly been guided by Derwent et al. 2010.

In the following definitions, O_i represents the ith observed value and M_i represents the ith modelled value for a total of n observations.

Fraction of predictions within a factor or two, *FAC2*

The fraction of modelled values within a factor of two of the observed values are the fraction of model predictions that satisfy:

$$0.5 \leq \frac{M_i}{O_i} \leq 2.0 \tag{1}$$

Mean bias, *MB*

The mean bias provides a good indication of the mean over or under estimate of predictions. Mean bias in the same units as the quantities being considered.

$$MB = \frac{1}{n} \sum_{i=1}^{N} M_i - O_i \tag{2}$$

Mean Gross Error, *MGE*

The mean gross error provides a good indication of the mean error regardless of whether it is an over or underestimate. Mean gross error is in the same units as the quantities being considered.

$$MGE = \frac{1}{n} \sum_{i=1}^{N} |M_i - O_i| \tag{3}$$

Normalised mean bias, *NMB*

The normalised mean bias is useful for comparing pollutants that cover different concentration scales and the mean bias is normalised by dividing by the observed concentration.

$$NMB = \frac{\sum_{i=1}^{n} M_i - O_i}{\sum_{i=1}^{n} O_i} \tag{4}$$

Normalised mean gross error, *NMGE*

The normalised mean gross error further ignores whether a prediction is an over or underestimate.

$$NMGE = \frac{\sum_{i=1}^{n} |M_i - O_i|}{\sum_{i=1}^{n} O_i} \tag{5}$$

Root mean squared error, *RMSE*

The *RMSE* is a commonly used statistic that provides a good overall measure of how close modelled values are to predicted values.

$$RMSE = \left(\frac{\sum_{i=1}^{n} (M_i - O_i)^2}{n} \right)^{1/2} \tag{6}$$

Correlation coefficient, *r*

The (Pearson) correlation coefficient is a measure of the strength of the linear relationship between two variables. If there is perfect linear relationship with positive slope between the two variables, $r = 1$. If there is a perfect linear relationship with negative slope between the two variables $r = -1$. A correlation coefficient of 0 means that there is no linear relationship between the variables.

$$r = \frac{1}{(n-1)} \sum_{i=1}^{n} \left(\frac{M_i - \overline{M}}{\sigma_M} \right) \left(\frac{O_i - \overline{O}}{\sigma_O} \right) \tag{7}$$

Coefficient of Efficiency, *COE*

The *Coefficient of Efficiency* based on Legates and McCabe (2012) and Legates and McCabe Jr (1999). There have been many suggestions for measuring model performance over the years, but the *COE* is a simple formulation which is easy to interpret.

A perfect model has a $COE = 1$. As noted by Legates and McCabe although the *COE* has no lower bound, a value of $COE = 0.0$ has a fundamental meaning. It implies that the model is no more able to predict the observed values than does the observed mean. Therefore, since the model can explain no more of the variation in the observed values than can the observed mean, such a model can have no predictive advantage.

For negative values of *COE*, the model is less effective than the observed mean in predicting the variation in the observations.

$$COE = 1.0 - \frac{\sum_{i=1}^{n} |M_i - O_i|}{\sum_{i=1}^{n} |O_i - \overline{O}|} \tag{8}$$

C Code used to produce outputs

This section contains all the code that produces the Figures and Tables. The data used in this report are stored in a dropbox folder owned by David Carslaw. Users of the data can contact David Carslaw at mailto:david.carslaw@kcl.ac.uk to request access. The data itself can then be downloaded to a suitable local location. Note that in order for the code to run, users would need to change the path of some of the file locations. There is also a ReadMe.txt file that should be read.

```
 ## ----LoadPackages----------------------------------------------------------
library(openair)
library(plyr)
library(ggplot2)
library(xtable)
library(reshape2)
## set ggplot plots to B&W
theme_set(theme_bw())
theme_update(strip.background = element_rect(colour = 'black'))
## make sure wd is correctly set
setwd("~/Projects/modelEvaluation/phase2/urban/")
## ----Loaddata, cache=FALSE-------------------------------------------------
load("./data/sourceApp.Rdata")
## catergorise as roadside or background
id <- grep("urban", nox.SA$site.type)
nox.SA$agg.type <- "roadside"
nox.SA$agg.type[id] <- "background"
id <- grep("urban", pm10.SA$site.type)
pm10.SA$agg.type <- "roadside"
pm10.SA$agg.type[id] <- "background"
id <- grep("urban", pm25.SA$site.type)
pm25.SA$agg.type <- "roadside"
pm25.SA$agg.type[id] <- "background"
## ----LoadHourly------------------------------------------------------------
## urban measurements
load("./data/urbanMeas.RData")
load("./data/hourlyPreds.RData")
## heathrow met and WRF data
load("./data/metData.RData")
met <- subset(met, select = c(date, ws, wd))
CERC.hourly$data <- "ADMS-Urban"
CERC.hourly <- subset(CERC.hourly, scenario == "Base")
CERC.hourly <- merge(CERC.hourly, met, by = "date")
CERC.hourly <- merge(CERC.hourly, urban.meas, by = c("code", "date"),
                     suffixes = c(".mod", ".obs"))
KCL.hourly$data <- "KCL-CMAQ-urban"
KCL.hourly$code <- as.character(KCL.hourly$code)
KCL.hourly <- subset(KCL.hourly, scenario == "Base")
KCL.hourly <- merge(KCL.hourly, wrfMet, by = "date")
KCL.hourly <- merge(KCL.hourly, urban.meas, by = c("code", "date"),
                     suffixes = c(".mod", ".obs"))
allData <- rbind.fill(CERC.hourly, KCL.hourly)
## calculate AQ indexes
labs <- c("Low.1", "Low.2", "Low.3", "Moderate.4", "Moderate.5", "Moderate.6", "High.7",
          "High.8", "High.9", "Very High.10")
allData <- ddply(allData, .(data), rollingMean, pollutant="O3.obs",
                 new.name = "O3.roll.obs")
allData <- ddply(allData, .(data), rollingMean, pollutant="O3.mod",
                 new.name = "O3.roll.mod")
allData <- transform(allData, index.O3 = cut(O3.roll.obs,
                             breaks = c(0, 34, 66, 100, 121, 141, 160,
                             188, 214, 240, 500), labels=labs),
                     index.NO2 = cut(NO2.obs, breaks = c(0, 67, 134, 200,
                                     268, 335, 400, 468, 535,
                                     600, 1000), labels = labs),
                     index.NO2.mod = cut(NO2.mod, breaks = c(0, 67, 134, 200,
                                         268, 335, 400, 468, 535,
                                     600, 1000), labels = labs))
## daily means for index for PM10/PM2.5
daily <- ddply(allData, .(data), timeAverage, avg.time = "day")
```

```
daily <- transform(daily,  index.PM10 = cut(PM10.obs, breaks = c(0, 17, 34, 50,
                                                    59, 67, 75, 84, 92, 100, 500),
                     labels = labs),
                     index.PM25 = cut(PM25.obs, breaks = c(0, 12, 24, 35, 42,
                                                47, 53, 59, 65, 70, 500),
                     labels = labs))
## calculate annual values for simple comparsions
annual <- timeAverage(urban.meas, "year")
## ----SANOx,results='asis',echo=FALSE-----------------------------------
tmp.nox <- ddply(subset(nox.SA, select = c(model, agg.type, variable, value)),
                 .(model, agg.type, variable), numcolwise(mean))
tmp.nox <- dcast(tmp.nox, ...~ model)
print(xtable(tmp.nox,
      caption = "Mean source contributions to \\nox concentrations by model
and major source category for 16 sites (\\ug).",
              label = "tab:SANOx",
              digits = 1),
      size = "footnotesize",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----SANOxPercent,results='asis',echo=FALSE----------------------------
m.nox <- ddply(tmp.nox, .(agg.type), function (x)
               100 * sweep(x[, 3:6], 2, colSums(x[, 3:6]), "/"))
m.nox <- data.frame(type = m.nox$agg.type, variable = tmp.nox$variable, m.nox[, 2:5])
print(xtable(m.nox,
      caption = "Mean source contributions to \\nox concentrations by
model and major source category for 16 sites (\\%).",
              label = "tab:SANOxPercent",
              digits =1),
      size = "footnotesize",
      booktabs = TRUE,
      include.rownames = FALSE,
      caption.placement = "top")
## ----NOxSA, h=10, w=14, out.width='1\\textwidth'-----------------------
ggplot(nox.SA, aes(model, fill=variable, weight = value)) +
    geom_bar(width = 0.5) +
    facet_wrap(~ site.name, scales = "free_y") +
    scale_fill_brewer(palette = "Set3") +
    geom_hline(aes(yintercept = nox), data = nox.meas, lty = 5) +
    ylab(quickText("nox (ug/m3)")) +
    theme(axis.text.x = element_text(angle= -90, hjust = 0))
## ----statsNOx----------------------------------------------------------
nox.annual <- ddply(nox.SA, .(site.name, model), numcolwise(sum), na.rm=TRUE)
nox.annual <- merge(nox.annual, annual, by.x = "site.name", by.y = "site")
## ----SAPm10,results='asis', echo=FALSE---------------------------------
tmp.pm10 <- ddply(subset(pm10.SA, select = c(model, agg.type, variable, value)),
                  .(model, agg.type, variable), numcolwise(mean))
tmp.pm10 <- dcast(tmp.pm10, ...~ model)
print(xtable(tmp.pm10,
      caption = "Mean source contributions to \\pmten concentrations
by model and major source category for 16 sites (\\ug).",
              label = "tab:SAPm10",
              digits = 1),
      size = "footnotesize",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----SAPm10Percent,results='asis',echo=FALSE---------------------------
m.pm10 <- ddply(tmp.pm10, .(agg.type), function (x)
               100 * sweep(x[, 3:6], 2, colSums(x[, 3:6]), "/"))
m.pm10 <- data.frame(type = m.pm10$agg.type, variable = tmp.pm10$variable, m.pm10[, 2:5])
print(xtable(m.pm10,
      caption = "Mean source contributions to \\pmten concentrations
by model and major source category for 16 sites (\\%).",
              label = "tab:SAPm10Percent",
              digits = 1),
      size = "footnotesize",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----Pm10SA, h=10, w=14, out.width='1\\textwidth'-----------------------
```

```
ggplot(pm10.SA, aes(model, fill=variable, weight = value)) +
    geom_bar(width = 0.5) +
    facet_wrap(~ site.name, scales = "free_y") +
    scale_fill_brewer(palette = "Set3") +
    geom_hline(aes(yintercept = PM10), data = pm10.meas, lty = 5) +
    ylab(quickText("pm10 (ug/m3)")) +
    theme(axis.text.x = element_text(angle= -90, hjust = 0))
## ----backSA, h=4, w=7, out.width='0.6\\textwidth'-----------------------
back.SA <- read.csv("./data/background_SA.csv",header=TRUE)
ggplot(back.SA, aes(model, fill = component, weight = value)) +
    geom_bar(width = 0.5) +
    scale_fill_brewer(palette="Set3") +
    ylab(quickText("concentration (ug/m3)"))
## ----SAPM25,results='asis',echo=FALSE-----------------------------------
tmp.pm25 <- ddply(subset(pm25.SA, select = c(model, agg.type, variable, value)),
                  .(model, agg.type, variable), numcolwise(mean))
tmp.pm25 <- dcast(tmp.pm25, ...~ model)
print(xtable(tmp.pm25,
      caption = "Mean source contributions to \\pmtwo concentrations
by model and major source category for 16 sites (\\ug).",
              label = "tab:SAPM25",
              digits = 1),
      size = "footnotesize",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----SAPM25Percent,results='asis',echo=FALSE----------------------------
m.pm25 <- ddply(tmp.pm25, .(agg.type), function (x)
                100 * sweep(x[, 3:6], 2, colSums(x[, 3:6]), "/"))
m.pm25 <- data.frame(type = m.pm25$agg.type, variable = tmp.pm25$variable, m.pm25[, 2:5])
print(xtable(m.pm25,
      caption = "Mean source contributions to \\pmtwo concentrations
by model and major source category for 16 sites (\\%).",
              label = "tab:SAPM25Percent",
              digits = 1),
      size = "footnotesize",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----Pm25SA, h=10, w=14, out.width='1\\textwidth'------------------------
ggplot(pm25.SA, aes(model, fill=variable, weight = value)) +
    geom_bar(width = 0.5) +
    facet_wrap(~ site.name, scales = "free_y") +
    scale_fill_brewer(palette = "Set3") +
    geom_hline(aes(yintercept = PM25), data = pm25.meas, lty = 5) +
    ylab(quickText("pm25 (ug/m3)")) +
    theme(axis.text.x = element_text(angle= -90, hjust = 0))
## ----LoadMapData--------------------------------------------------------
load("./data/mapData.RData")
load("./data/grid100.RData")
## ----no2Map,h=7, w=9, out.width='1\\textwidth'---------------------------
scatterPlot(map.data, x="easting", y="northing", z = "no2",
            method ="level", x.inc = 1000,
            y.inc=1000, limits = c(0, 50), type = "model",
            main = NULL)
## ----no2MapHiRes,h=3.5, w=7, out.width='1\\textwidth'--------------------
scatterPlot(grid100, x="easting", y="northing",
            z="no2", method ="level", x.inc = 100,
            y.inc=100, limits = c(0, 50),
            type = "model", main = NULL)
## ----NOxTransect, h=5,w=5,fig.show='hold',out.width='0.49\\textwidth'----
transect <- subset(map.data, northing == 182000)
scatterPlot(transect, x = "easting", y = "nox",  group = "model", plot.type = "S",
            lwd = 3, ylab = "nox (ug/m3)", key.position = "top")
scatterPlot(transect, x = "easting", y = "no2",  group = "model", plot.type = "S",
            lwd = 3, ylab = "no2 (ug/m3)", key.position = "top")
## ----NO2exceedcum, h=5,w=5,fig.show='hold',out.width='0.45\\textwidth'----
no2.data <- ddply(map.data, .(model), transform, nox = sort(nox, na.last = TRUE))
no2.data <- ddply(no2.data, .(model), transform, area = 100 * seq(1, length(nox)) / length(nox))
scatterPlot(no2.data, x="nox", y="area", plot.type="l", lwd=4,
            group = "model", xlim = c(10, 100), col = "hue",
            xlab = "nox (ug/m3)", ylab = "Cumulative area (%)", key.position = "top",
```

```
            key.columns = 2)
no2.data <- ddply(map.data, .(model), transform, no2 = sort(no2, na.last = TRUE))
no2.data <- ddply(no2.data, .(model), transform, area = 100 * seq(1, length(no2)) / length(no2))
scatterPlot(no2.data, x="no2", y="area", plot.type="l", lwd=4,
            group = "model", xlim = c(10, 60), col = "hue",
            xlab = "no2 (ug/m3)", ylab = "Cumulative area (%)", key.position = "top",
            key.columns = 2)
no2.data <- ddply(map.data, .(model), transform, ratio = sort(ratio, na.last = TRUE))
no2.data <- ddply(no2.data, .(model), transform, area = 100 * seq(1, length(ratio)) / length(ratio))
scatterPlot(no2.data, x="ratio", y="area", plot.type="l", lwd=4,
            group = "model", xlim = c(0.4, 0.8), col = "hue",
            xlab = "no2/nox ratio", ylab = "Cumulative area (%)", key.position = "top",
            key.columns = 2)
## ----no2noxMap,h=7, w=9, out.width='1\\textwidth'-----------------------
scatterPlot(map.data, x="easting", y ="northing", z="ratio", method ="level",
            x.inc = 1000, y.inc=1000, type = "model", main = NULL,
            limits = c(0.35, 0.9))
## ----pm10Map,h=7, w=9, out.width='1\\textwidth'-------------------------
scatterPlot(map.data, x="easting", y="northing", z="pm10", method ="level", x.inc = 1000,
            y.inc=1000, limits = c(15, 30), type = "model", main = NULL)
## ----PMTransect, h=5,w=5,fig.show='hold',out.width='0.49\\textwidth'-----
transectPM <- subset(map.data, northing == 182000)
scatterPlot(transect, x = "easting", y = "pm10",  group = "model", plot.type = "S",
            lwd = 3, ylab = "pm10 (ug/m3)", key.position = "top")
scatterPlot(transect, x = "easting", y = "pm25",  group = "model", plot.type = "S",
            lwd = 3, ylab = "pm25 (ug/m3)", key.position = "top")
## ----PM10exceedcum, h=4, w=6, out.width='0.6\\textwidth'-----------------
## calculate areas
pm10.data <- arrange(map.data, model, pm10)
pm10.data <- ddply(pm10.data, .(model), transform, area = 100 * seq(1, length(pm10)) / length(pm10))
scatterPlot(pm10.data, x="pm10", y="area", plot.type="l", lwd=4,
            group = "model", xlim = c(15, 25),
            xlab = "pm10 (ug/m3)", ylab = "Cumulative area (%)")
## ----pm25Map,h=7, w=9, out.width='1\\textwidth'-------------------------
scatterPlot(map.data, x="easting", y="northing", z="pm25", method ="level", x.inc = 1000,
            y.inc=1000, limits = c(8, 18), type = "model", main = NULL)
## ----PM25exceedcum, h=4, w=6, out.width='0.6\\textwidth'-----------------
pm25.data <- arrange(map.data, model, pm25)
pm25.data <- ddply(pm25.data, .(model), transform, area = 100 * seq(1, length(pm25)) / length(pm25))
scatterPlot(pm25.data, x="pm25", y="area", plot.type="l", lwd=4,
            group = "model", xlim = c(9, 18),
            xlab = "pm25 (ug/m3)", ylab = "Cumulative area (%)")
## ----loadScenarios------------------------------------------------------
load("./data/scenarios.RData")
## ----scenResults,results='asis',echo=FALSE------------------------------
print(xtable(ddply(subset(urban.scen, select = -Base), .(pollutant, model),
                   numcolwise(mean), na.rm=T),
      caption = "Effect of the different scenarios by pollutant and model
(\\ug). The base case predictions have been subtracted from each scenario
to show the nett effect on concentrations.",
             label = "tab:scenResults",
             digits = 1),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----NO2ScenSite, h=4, w=6, out.width='0.7\\textwidth'-------------------
## sort data frame by scenario 1
tmpNO2 <- openair:::sortDataFrame(subset(urban.scen, pollutant == "NO2"), key = c("model", "S1"))
## match sites - PCM has some missing...
scen.sites <- subset(urban.scen, model == "PCM", select = c(code, S1))
scen.sites <- unique(na.omit(scen.sites)$code)
tmpNO2 <- subset(tmpNO2, code %in% scen.sites)
## create index for plotting
tmpNO2 <- ddply(tmpNO2, .(model, pollutant), transform, index = rev(seq_along(model)))
## for NOx
tmpNOx <- openair:::sortDataFrame(subset(urban.scen, pollutant == "NOx"), key = c("model", "S1"))
## match sites - PCM has some missing...
scen.sites <- subset(urban.scen, model == "PCM", select = c(code, S1))
scen.sites <- unique(na.omit(scen.sites)$code)
tmpNOx <- subset(tmpNOx, code %in% scen.sites)
## create index for plotting
```

```
tmpNOx <- ddply(tmpNOx, .(model, pollutant), transform, index = rev(seq_along(model)))
## NOx and NO2
tmpDiff <- merge(subset(tmpNO2, select = c(code, model, S1)), subset(tmpNOx, select = c(code, model, S1)),
                 by = c("code", "model"))
## plot
scatterPlot(tmpNO2, x = "index", y = "S1", group = "model", pch = NA, plot.type = "l",
            lwd = 3, cols = "hue", xlab = "site index", ylab = "delta no2 (ug/m3)")
## ----NO2NOxScenSite, h=4, w=5.5, out.width='0.7\\textwidth'--------------
scatterPlot(tmpDiff, x = "S1.y", y = "S1.x", group = "model", pch = c(15:17,3,2),
            cols = "hue",
            xlab = "delta nox (ug/m3)", ylab = "delta no2 (ug/m3)")
## ----modStatsNOxhourly,results='asis',echo=FALSE-------------------------
modNOx <- modStats(allData, obs="NOx.obs", mod = "NOx.mod", type = "data", rank.name="data")
modNOx <- melt(modNOx, id.vars = "data")
modNOx <- dcast(modNOx, ... ~ data)
print(xtable(modNOx,
      caption = "Performance of \\adms and \\cmaq for \\nox concentrations.",
             label = "tab:modStatsNOxhourly",
             digits = 2),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----modStatsNO2hourly,results='asis',echo=FALSE-------------------------
modNO2 <- modStats(allData, obs="NO2.obs", mod = "NO2.mod", type = "data", rank.name="data")
modNO2 <- melt(modNO2, id.vars = "data")
modNO2 <- dcast(modNO2, ... ~ data)
print(xtable(modNO2,
      caption = "Performance of \\adms and \\cmaq for \\notwo concentrations.",
             label = "tab:modStatsNO2hourly",
             digits = 2),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----modStatsO3hourly,results='asis',echo=FALSE--------------------------
modO3 <- modStats(allData, obs="O3.obs", mod = "O3.mod", type = "data", rank.name="data")
modO3 <- melt(modO3, id.vars = "data")
modO3 <- dcast(modO3, ... ~ data)
print(xtable(modO3,
      caption = "Performance of \\adms and \\cmaq for \\ozone concentrations.",
             label = "tab:modStatsO3hourly",
             digits = 2),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----modStatsPM10hourly,results='asis',echo=FALSE------------------------
modPM10 <- modStats(allData, obs="PM10.obs", mod = "PM10.mod", type = "data", rank.name="data")
modPM10 <- melt(modPM10, id.vars = "data")
modPM10 <- dcast(modPM10, ... ~ data)
print(xtable(modPM10,
      caption = "Performance of \\adms and \\cmaq for \\pmten concentrations.",
             label = "tab:modStatsPM10hourly",
             digits = 2),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----modStatsPM25hourly,results='asis',echo=FALSE------------------------
modPM25 <- modStats(allData, obs="PM25.obs", mod = "PM25.mod", type = "data", rank.name="data")
modPM25 <- melt(modPM25, id.vars = "data")
modPM25 <- dcast(modPM25, ... ~ data)
print(xtable(modPM25,
      caption = "Performance of \\adms and \\cmaq for \\pmtwo concentrations.",
             label = "tab:modStatsPM25hourly",
             digits = 2),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----prepMetrics---------------------------------------------------------
```

```
## calculate common AQ stats and compare with observed values
obsStats <- aqStats(subset(allData, data == "ADMS-Urban"), pollut="NO2.obs")
obsStats$data <- "observed"
admsStats <- aqStats(subset(allData, data == "ADMS-Urban"), pollut="NO2.mod")
admsStats$data <- "ADMS-Urban"
kclStats <- aqStats(subset(allData, data == "KCL-CMAQ-urban"), pollut="NO2.mod")
kclStats$data <- "KCL-CMAQ-urban"
no2Stats <- rbind.fill(admsStats, kclStats)
no2Stats <- merge(no2Stats, obsStats, by = c("site", "pollutant", "year"),
                  suffixes = c(".mod", ".obs"))
## calculate common AQ stats and compare with observed values
obsStats <- aqStats(subset(allData, data == "ADMS-Urban"), pollut="O3.obs")
obsStats$data <- "observed"
admsStats <- aqStats(subset(allData, data == "ADMS-Urban"), pollut="O3.mod")
admsStats$data <- "ADMS-Urban"
kclStats <- aqStats(subset(allData, data == "KCL-CMAQ-urban"), pollut="O3.mod")
kclStats$data <- "KCL-CMAQ-urban"
o3Stats <- rbind.fill(admsStats, kclStats)
o3Stats <- merge(o3Stats, obsStats, by = c("site", "pollutant", "year"),
                  suffixes = c(".mod", ".obs"))
obsStats <- aqStats(subset(allData, data == "ADMS-Urban"), pollut="PM10.obs")
obsStats$data <- "observed"
admsStats <- aqStats(subset(allData, data == "ADMS-Urban"), pollut="PM10.mod")
admsStats$data <- "ADMS-Urban"
kclStats <- aqStats(subset(allData, data == "KCL-CMAQ-urban"), pollut="PM10.mod")
kclStats$data <- "KCL-CMAQ-urban"
pm10Stats <- rbind.fill(admsStats, kclStats)
pm10Stats <- merge(pm10Stats, obsStats, by = c("site", "pollutant", "year"),
                  suffixes = c(".mod", ".obs"))
## ----NO2Stats, h=4.5,w=4.5,fig.show='hold',out.width='0.32\\textwidth'----
scatterPlot(no2Stats, x = "mean.mod",  y = "mean.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, xlim = c(0, 220),
            ylim = c(0, 220), key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled no2 (ug/m3)",
            ylab = "observed no2 (ug/m3)")
scatterPlot(no2Stats, x = "max.daily.mod",  y= "max.daily.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, xlim = c(0, 250),
            ylim = c(0, 250), key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled no2 (ug/m3)",
            ylab = "observed no2 (ug/m3)")
scatterPlot(no2Stats, x = "hours.gt.200.mod",  y = "hours.gt.200.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled hours no2 > 200 ug/m3",
            ylab = "observed hours no2 > 200 ug/m3", xlim=c(1, 5000), ylim=c(1, 5000))
## ----O3Stats, h=4.5,w=4.5,fig.show='hold',out.width='0.32\\textwidth'----
scatterPlot(o3Stats, x = "mean.mod",  y= "mean.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, xlim = c(0, 50),
            ylim = c(0, 50), key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled o3 (ug/m3)",
            ylab = "observed o3 (ug/m3)")
scatterPlot(no2Stats, x = "max.rolling.8.mod",  y= "max.rolling.8.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, xlim = c(0, 250),
            ylim = c(0, 250), key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled o3 (ug/m3)",
            ylab = "observed o3 (ug/m3)")
scatterPlot(o3Stats, x = "roll.8.O3.gt.100.mod",  y= "roll.8.O3.gt.100.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled days",
            ylab = "observed days", xlim=c(1, 40), ylim=c(1, 40))
## ----PM10Stats, h=4.5,w=4.5,fig.show='hold',out.width='0.32\\textwidth'----
scatterPlot(pm10Stats, x = "mean.mod",  y = "mean.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, xlim = c(0, 42),
            ylim = c(0, 42), key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled pm10 (ug/m3)",
            ylab = "observed pm10 (ug/m3)")
scatterPlot(pm10Stats, x = "days.gt.50.mod",  y = "days.gt.50.obs", group = "data.mod",
            pch = c(16, 8), mod.line=T, cex = 1.5, xlim = c(0, 70),
            ylim = c(0, 70), key.position = "top", key.columns = 2,
            key.title = "model", xlab = "modelled pm10 (ug/m3)",
            ylab = "observed pm10 (ug/m3)")
## ----polarKC1NOx, h=4, w=8, out.width='1\\textwidth'--------------------
tmp <- subset(allData, code == "KC1")
```

```
## add measured for comparison
tmp1 <- subset(tmp, data == "ADMS-Urban")
tmp1 <- transform(tmp1, NOx.mod = NOx.obs, NO2.mod = NO2.obs,
                  PM10.mod = PM10.obs, PM25.mod = PM25.obs)
tmp1$data <- "MEASURED"
tmp <- rbind(tmp, tmp1)
polarPlot(subset(tmp, code == "KC1"), pollutant="NOx.mod", type = "data",
          layout = c(3, 1), limits = c(0, 150))
## ----SANOxKC1,results='asis',echo=FALSE--------------------------------
KC.SA <- ddply(subset(nox.SA, select = c(model, variable, value), site.code =="KC1"),
               .(model, variable), numcolwise(mean))
KC.SA <- dcast(KC.SA,... ~ model)
print(xtable(KC.SA,
      caption = "Mean source contributions to \\nox concentrations by model and
major source category for the North Kensignton site (\\ug).",
             label = "tab:SANOxKC1",
             digits = 1),
      size = "small",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----polarBL0NOx, h=4, w=8, out.width='1\\textwidth'--------------------
tmp <- subset(allData, code== "BL0")
## add measured for comparison
tmp1 <- subset(tmp, data =="ADMS-Urban")
tmp1 <- transform(tmp1, NOx.mod = NOx.obs, NO2.mod = NO2.obs,
                  PM10.mod = PM10.obs, PM25.mod = PM25.obs)
tmp1$data <- "MEASURED"
tmp <- rbind(tmp, tmp1)
polarPlot(subset(tmp, code == "BL0"), pollutant="NOx.mod",
          type = "data", layout = c(3, 1))
## ----polarMY1NOx, h=4, w=8, out.width='1\\textwidth'--------------------
tmp <- subset(allData, code== "MY1")
## add measured for comparison
tmp1 <- subset(tmp, data =="ADMS-Urban")
tmp1 <- transform(tmp1, NOx.mod = NOx.obs, NO2.mod = NO2.obs,
                  PM10.mod = PM10.obs, PM25.mod = PM25.obs)
tmp1$data <- "MEASURED"
tmp <- rbind(tmp, tmp1)
polarPlot(subset(tmp, code == "MY1"), pollutant="NOx.mod",
          type = "data", layout = c(3, 1))
## ----MY1NOxscatt, h=4, w=10, out.width='0.8\\textwidth'-----------------
scatterPlot(subset(allData, code == "MY1"), x = "NOx.obs", y = "NOx.mod", mod.line = TRUE,
            type = "data", method = "hexbin", col = "jet",
            xlab = "observed nox (ug/m3)",
            ylab = "modelled nox (ug/m3)", xbin = 40,
            log.x = TRUE, log.y = TRUE, aspect = 1,
            trans = function(x) log(x), inv = function(x) exp(x))
## ----condQuantMY1ws, h=7, w=15,fig.keep='last',out.width='1\\textwidth'----
tmp <- subset(allData, code== "MY1")
## add measured for comparison
tmp1 <- subset(tmp, data =="ADMS-Urban")
tmp1 <- transform(tmp1, NOx.mod = NOx.obs, NO2.mod = NO2.obs,
                  PM10.mod = PM10.obs, PM25.mod = PM25.obs)
tmp1$data <- "MEASURED"
tmp <- rbind(tmp, tmp1)
conditionalEval(tmp, mod = "NOx.mod", obs = "NOx.obs",
                type = "data", statistic = "ws", col.var = "Set3")
## ----condQuantMY1Sector, h=7,w=15,fig.keep='last',out.width='1\\textwidth'----
tmp$sector <- "south"
id <- which(tmp$wd <=90 | tmp$wd >=270)
tmp$sector[id] <- "north"
conditionalEval(tmp, mod = "NOx.mod", obs= "NOx.obs",
                type = "data", statistic = "sector", col.var = "Set3")
## ----polarMY1NO2, h=4, w=8, out.width='1\\textwidth'--------------------
tmp <- subset(allData, code== "MY1")
## add measured for comparison
tmp1 <- subset(tmp, data =="ADMS-Urban")
tmp1 <- transform(tmp1, NOx.mod = NOx.obs, NO2.mod = NO2.obs,
                  PM10.mod = PM10.obs, PM25.mod = PM25.obs)
tmp1$data <- "MEASURED"
tmp <- rbind(tmp, tmp1)
```

```
polarPlot(subset(tmp, code == "MY1"), pollutant="NO2.mod", stati="cpf",
          percentile=94, type = "data", layout = c(3, 1))
## ----importOSPM, results='hide'-----------------------------------------
ospm <- import("./data/OSPM_Validation_Tools_sample.csv", wd = "Wind_dir",
               ws = "u_mast")
## ----OSPMpolar,w=7,h=4,results='hide', out.width='0.7\\textwidth'--------
polarPlot(ospm, pollutant=c("cNOX_obs_2", "cNOX_mod_2"), layout = c(2, 1),
          limits = c(0, 150))
## ----prepHourlyNO2------------------------------------------------------
makeNOxNO2 <- function(thecode) {
    thedata <- subset(allData, code == thecode, select = c(site, code, data, NOx.obs,
                                                            NOx.mod, NO2.obs, NO2.mod))
    ## add measured at bottom to plot
    tmp <- subset(thedata, data == "ADMS-Urban")
    tmp$data <- "measured"
    tmp <- transform(tmp, NOx.mod = NOx.obs, NO2.mod = NO2.obs)
    thedata <- rbind(thedata, tmp)
    thedata <- subset(thedata, select = -c(NOx.obs, NO2.obs))
    ## round nox to interval
    thedata$nox <- round_any(thedata$NOx.mod, 5)
    thedata <- aggregate(thedata[, c("NOx.mod", "NO2.mod")],
                         thedata[, c("site", "code", "data", "nox")], mean, na.rm = TRUE)
    thedata <- transform(thedata, ratio = NO2.mod / NOx.mod)
    thedata
    }
## sites to consider
sites <- unique(allData$code)
nox.no2 <- ldply(sites, makeNOxNO2)
stats.ratio <- merge(subset(nox.no2, data == "measured"),
                     subset(nox.no2, data != "measured"), by = c("site", "nox"))
## get site info
meta <- importMeta(source="kcl")
stats.ratio <- merge(stats.ratio, meta,  by = "site")
## ----modStatsRatio,results='asis',echo=FALSE----------------------------
statsRatio <- modStats(stats.ratio, obs = "ratio.x", mod = "ratio.y",
                       type = c("site.type", "data.y"), rank = "data.y")
names(statsRatio)[1:2] <- c("site type", "model")
statsRatio[, 1] <-  abbreviate(statsRatio[, 1])
print(xtable(statsRatio,
      caption = "Performance of \\adms and \\cmaq for \\notwo:\\nox ratio by site type and model.
Note for site type Krbs = kerbside, Rdsd = roadside, Sbrb = suburban and UrbB = urban background. ",
             label = "tab:modStatsRatio",
             digits = 2),
      size = "scriptsize",
      booktabs = TRUE,
      include.rownames=FALSE,
      caption.placement = "top")
## ----NOxNO2Relation, h=10, w=10,fig.show='hold',results='hide', out.width='1\\textwidth'----
scatterPlot(subset(nox.no2, nox >0),
            x = "nox", y = "ratio", type = "site",
            group = "data", plot.type = "l",
            col = c("magenta", "cyan", "grey40"), lwd = 3,
            ylab = "no2:nox ratio",  xlab = "nox (ug/m3)",
            key.position = "top", key.columns = 3,
            log.x = TRUE)
## ----AQindexMBNO2,w=6,h=5,fig.show='hold', out.width='0.49\\textwidth'----
res <- modStats(subset(allData, code != "LB4"), obs = "NO2.obs", mod = "NO2.mod",
                type = c("index.NO2", "data"), rank.name="data")
scatterPlot(res, x = "index.NO2", y = "MB", group = "data", pch = 16, cex=2,
            ref.y=0, ylab = "MB (ug/m3)")
scatterPlot(res, x = "index.NO2", y = "NMB", group = "data", pch = 16, cex=2, ref.y=0)
## ----AQindexMBO3,w=6,h=5,fig.show='hold', out.width='0.49\\textwidth'----
res <- modStats(subset(allData, code != "LB4"), obs = "O3.roll.obs", mod = "O3.roll.mod",
                type = c("index.O3", "data"), rank.name="data")
scatterPlot(res, x = "index.O3", y = "MB", group = "data", pch = 16, cex=2, ref.y=0, ylab = "MB (ug/m3)")
scatterPlot(res, x = "index.O3", y = "NMB", group = "data", pch = 16, cex=2, ref.y=0)
## ----AQindexMBPM10,w=6,h=5,fig.show='hold', out.width='0.49\\textwidth'----
res <- modStats(subset(daily, site != "Lambeth - Brixton Road"), obs = "PM10.obs", mod = "PM10.mod",
                type = c("index.PM10", "data"), rank.name="data")
scatterPlot(res, x = "index.PM10", y = "MB", group = "data", pch = 16, cex=2, ref.y=0, ylab = "MB (ug/m3)")
scatterPlot(res, x = "index.PM10", y = "NMB", group = "data", pch = 16, cex=2, ref.y=0)
## ----AQindexMBPM25,w=6,h=5,fig.show='hold', out.width='0.49\\textwidth'----
```

```
res <- modStats(subset(daily, site != "Lambeth - Brixton Road"), obs = "PM25.obs", mod = "PM25.mod",
                type = c("index.PM25", "data"), rank.name="data")
scatterPlot(res, x = "index.PM25", y = "MB", group = "data", pch = 16, cex=2, ref.y=0, ylab = "MB (ug/m3)")
scatterPlot(res, x = "index.PM25", y = "NMB", group = "data", pch = 16, cex=2, ref.y=0)
## ----NOxSAScaled, h=10, w=14, out.width='1\\textwidth'-------------------
## proportion plot
ggplot(nox.SA, aes(model, fill=variable, weight = value)) +
    geom_bar(width = 0.5, position="fill") +
    facet_wrap(~ site.name) +
    scale_fill_brewer(palette="Set3") +
    theme(axis.text.x = element_text(angle= -90, hjust = 0))
## ----Pm10SAScaled, h=10, w=14, out.width='1\\textwidth'-------------------
## proportion plot
ggplot(pm10.SA, aes(model, fill=variable, weight = value)) +
    geom_bar(width = 0.5, position="fill") +
    facet_wrap(~ site.name) +
    scale_fill_brewer(palette="Set3") +
    theme(axis.text.x = element_text(angle= -90, hjust = 0))
## ----Pm25SAScaled, h=10, w=14, out.width='1\\textwidth'-------------------
## proportion plot
ggplot(pm25.SA, aes(model, fill=variable, weight = value)) +
    geom_bar(width = 0.5, position="fill") +
    facet_wrap(~ site.name) +
    scale_fill_brewer(palette="Set3") +
    theme(axis.text.x = element_text(angle= -90, hjust = 0))
```